AL-TR-1992-0002

AD-A252 344

DEVELOPMENT OF A COMPOSITE MEASURE FOR PREDICTING ENGAGEMENT OUTCOME DURING AIR COMBAT MANEUVERING

ARMSTRONG LABORATORY

Wayne L. Waag

HUMAN RESOURCES DIRECTORATE
AIRCREW TRAINING RESEARCH DIVISION
Williams Air Force Base, AZ 85240-6457

William B. Raspotnik
Jeffrey L. Leeds

Logicon, Incorporated
Tactical and Training Systems Division
4010 Sorrento Valley Boulevard
P.O. Box 85158
San Diego, CA 92138-5158

May 1992

Final Technical Report for Period April 1987 - December 1991

AIR FORCE SYSTEMS COMMAND
BROOKS AIR FORCE BASE, TEXAS 78235-5000

NOTICES

When Government drawings, specifications, or other data are used for any purpose other than in connection with a definitely Government-related procurement, the United States Government incurs no responsibility or any obligation whatsoever. The fact that the Government may have formulated or in any way supplied the said drawings, specifications, or other data, is not to be regarded by implication, or otherwise in any manner construed, as licensing the holder, or any other person or corporation; or as conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

The Office of Public Affairs has reviewed this report, and it is releasable to the National Technical Information Service, where it will be available to the general public, including foreign nationals.

This report has been reviewed and is approved for publication.

Wayne L. Waag

WAYNE L. WAAG
Contract Monitor

Dee H. Andrews

DEE H. ANDREWS, Technical Director
Aircrew Training Research Division

W. M. KORNOVICH, JR., Lt Col, USAF
Chief, Aircrew Training Research Division

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)

2. REPORT DATE
May 1992

3. REPORT TYPE AND DATES COVERED
Final – April 1987 – December 1991

4. TITLE AND SUBTITLE
Development of a Composite Measure for Predicting Engagement Outcome During Air Combat Maneuvering

5. FUNDING NUMBERS
C – F33615-86-C-0012
PE – 62205F
PR – 1123
TA – 35
WU – 06

6. AUTHOR(S)
Wayne L. Waag
William B. Raspotnik
Jeffrey L. Leeds

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Logicon, Incorporated
Tactical and Training Systems Division
4010 Sorrento Valley Blvd
P.O. Box 85158
San Diego, CA 92138-5158

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAMES(S) AND ADDRESS(ES)
Armstrong Laboratory
Human Resources Directorate
Aircrew Training Research Division
Williams Air Force Base, AZ 85240-6457

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
AL-TR-1992-0002

11. SUPPLEMENTARY NOTES

Armstrong Laboratory Technical Monitor: Dr. Wayne L. Waag, (602) 988-6561, Ext. 115.

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

This research was to develop a composite performance measure for predicting engagement outcome during air combat maneuvering (ACM). Data were collected in the Simulator for Air-to-Air Combat (SAAC) located at Luke AFB, AZ. Each of 125 U.S. Air Force pilots current in the F-15 or F-16 fighter aircraft flew 8 SAAC 1v1 engagements. Experimental factors included Aircraft type, Start Position, Opponent type, and Order of engagement. Engagement outcomes were classified as win, loss, or draw. Ten candidate measures were evaluated reflecting basic aircraft state parameters and positional advantage measures indicating relative offensiveness and defensiveness. Objectives were to (1) determine effects of experimental factors on engagement outcome; (2) determine statistical relationship of candidate measures to experimental factors; (3) determine relative importance of measurement categories for prediction of engagement outcome; and (4) produce a composite measure of performance from all candidate measures that maximizes the prediction of engagement outcome. Results concluded that a composite measure of ACM performance from a linear combination of aircraft state and positional advantage measures for use in transfer of training evaluations is feasible. This investigation demonstrated that engagement outcome can be predicted from measures available from instrumented range systems such as the Air Combat Maneuvering Instrumentation. Implications for future research are discussed with recommendations for analyses of this engagement database.

14. SUBJECT TERMS
Air combat
Air combat maneuvering
Flight simulation
Flight simulators
Flight training measures
Performance measurement
Tactical training

15. NUMBER OF PAGES
70

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18
298-102

CONTENTS

CONTENTS (Continued)

List of Figures

List of Tables

PREFACE

This research was conducted by the Human Resources Directorate, Aircrew Training Research Division (AL/HRA), Williams AFB, Arizona, under Work Unit 1123-35-06, Validation and Refinement of Techniques for Air Combat Performance Assessment. The effort was supported by Contract No. F33615-86-C-0012 with the Tactical and Training Systems Division of Logicon, Inc. Contract monitor is Dr Wayne L. Waag, AL/HRAT.

The authors wish to thank Dr Dave Hubbard, University of Dayton Research Institute (UDRI) for his assistance in data analysis. Appreciation is also expressed to Dr Don Lyon, also of UDRI, and Dr Elizabeth Martin, AL/HRA, for their review of this document. And finally, special thanks is given to Lt Col Dave Greschke, the TAC ACES Commander at Luke AFB, Arizona, who made possible data collection in the Simulator for Air-to-Air Combat. Without his personal support and the efforts of his staff, this research investigation would not have been possible.

DEVELOPMENT OF A COMPOSITE MEASURE FOR PREDICTING ENGAGEMENT OUTCOME DURING AIR COMBAT MANEUVERING

SUMMARY

The purpose of this research was to develop a composite measure of performance for predicting engagement outcome during air combat maneuvering (ACM). The data for this study were collected in the Simulator for Air-to-Air Combat (SAAC) located at Luke AFB, AZ. Each of 125 Air Force pilots who were current in the F-15 or F-16 fighter aircraft flew 8 one-versus-one (1V1) engagements in the SAAC for a total sample of 1,000 engagements. Experimental factors included Aircraft type (F-15 vs. F-16), Start Position (head-on vs. line abreast), Opponent type (human vs. computer-driven adversary) and Order of engagement presentation. Engagement outcomes were classified as either a win, loss, or draw. A total of 10 candidate measures were evaluated that reflected two categories: (1) basic aircraft state parameters such as altitude, airspeed, etc.; and (2) positional advantage measures indicating relative offensiveness and defensiveness.

Specific study objectives were to (1) determine the effects of the experimental factors on engagement outcome; (2) determine the statistical relationship of each candidate measure to each experimental factor; (3) determine the relative importance of each measurement category (i.e., positional advantage and aircraft state) for the prediction of engagement outcome; and (4) produce a composite measure of performance from all candidate measures that maximizes the prediction of engagement outcome. Results pertinent to these four research objectives were as follows:

1. Of the four experimental factors investigated, only Opponent significantly affected Outcome. This effect was quite large and is considered to be the most important factor within the engagement database that was gathered. Of the 500 engagements flown against each opponent, subjects defeated the Adaptive Maneuvering Logic (AML) 389 times and, in turn, were defeated by the Human opponent 362 times. The other factors (Aircraft, Start Position, and Order) produced no significant effects.

2. A large number of significant effects were obtained between the experimental factors and the individual candidate measures of performance. The one exception was Order, which produced no significant effects. Aircraft type produced a number of effects upon state measures which were due to differences in the flight control systems of the F-15 and F-16. Start Position also produced effects upon both state and positional advantage measures which were easily explained as a function of the initial set-up conditions. Opponent also produced effects on both types of measures. Differences in the state measures were primarily attributed to "style" differences between the AML and Human opponents, while differences in the position advantage measures were the result of the Human being a much more difficult opponent. And finally,

differences in both state and positional advantage measures occurred as a function of engagement Outcome. Strong linear relations were obtained between the "offensive" measures and Outcome, while the "defensive" measure was most discriminative of Losses. These findings provide support for the validity of these measures as predictors of engagement outcome.

3. Of the two categories of measures, aircraft state and positional advantage, the results clearly indicated that the positional advantage type was most highly related to engagement Outcome.

4. Composite measures were successfully developed for each Opponent for both a Win/Lose and Win/Draw/Lose criterion. In general, it was found that classification accuracy was highest for the Win/Lose criterion. For the Win/Lose/Draw criterion, classification accuracy for both the AML and Human Opponents was roughly 70% while accuracy for the Win/Lose criterion was 84% for the Human and 97% for the AML. Classification functions developed for one opponent (either the AML or Human) did not generalize to the other opponent, despite the fact that the same measures were selected for inclusions into the prediction equation. These findings suggest that the measures, themselves, are valid across opponents, but that development of an optimal set of weights is highly dependent upon the characteristics of the opponent.

It is concluded that development of a composite measure of ACM performance from a linear combination of aircraft state and positional advantage measures for use in transfer of training evaluations is feasible. The present investigation has demonstrated that engagement outcome can be predicted from measures available from the instrumented range systems such as the Air Combat Maneuvering Instrumentation (ACMI). Implications for future research are also discussed with specific recommendations for future analyses of this engagement database.

INTRODUCTION

This report documents exploratory efforts to develop a set of objective performance measures that can be used for evaluating the effectiveness of air combat training. The long-range goal of this work is to develop a reliable and valid measure of maneuvering performance that can be used in transfer of training evaluations. Specifically, the purpose of this initial investigation was to evaluate a set of candidate performance measures in terms of their ability to predict the outcome of simulated 1V1 air combat engagements.

Background

One must have a well-defined purpose in mind prior to the initiation of any measurement development activity. In an analysis of performance measurement requirements to support tactical aircrew training, Waag, Pierce, & Fessler (1987)

identified four broad categories of measurement use. These included performance monitoring, proficiency evaluation, training management, and training evaluation. In their analysis of the adequacy of existing capabilities, they concluded that there were shortfalls in two of these areas, namely performance monitoring and training evaluation. The problem with performance monitoring was the lack of information, especially within the in-flight environment, that would enable the proper diagnosis of performance deficiencies and hence suggest remedial action.

However, the area considered to be most deficient was training evaluation. Since the function of training is to prepare the aircrew for performance in the operational environment, there is a need for information that is reflective of that ability. An additional requirement for training evaluation has to do with the assessment of the impact that modifications to programs have on aircrew proficiency. Training programs and their individual components such as flight simulators, must be evaluated in terms of how well they prepare the aircrews for mission performance. Any improvements should be reflected by better individual aircrew performance at the operational level. Clearly, the evaluation of the effectiveness of such training programs and their individual components requires valid performance measurement information.

Within the arena of air combat training, one issue of interest to the Air Force is the potential contribution of flight simulation. In particular, what is the value of air combat simulation? The value of all ground-based training, and simulation in particular, can best be evaluated in terms of its impact upon subsequent performance in the in-flight environment. While the design of such evaluations is fairly straightforward conceptually, their successful implementation in a practical sense is quite difficult. For that reason, there have been very few studies that have evaluated the effectiveness of air combat simulation, and even fewer that have produced convincing results. In fact, the only studies reported were done in the late 70s and early 80s. In a study for the U.S. Navy, Payne et al. (1976) reported positive transfer for basic fighter maneuvers using the Northrop simulation facility. Two other transfer studies were accomplished in the U.S. Air Force's Simulator for Air-to-Air Combat located at Luke AFB. One showed no significant effect (Pohlmann & Reed, 1978) for teaching basic fighter maneuvers. The other showed a small effect (Jenkins, 1982) of the training on subsequent aircrew performance during a Fighter Weapons Instructor Course.

Although the problems of conducting transfer of training evaluations in the real world are many, it is clear that one of the major obstacles, especially in the air combat arena, is the lack of reliable and valid measures of performance. On the surface, it may appear to be a simple task in that engagements can easily be characterized by their outcomes (win/lose/draw). This is true assuming that aircrews can be observed over a fairly large number of engagements. However, another real-world constraint is that the amount of data available for evaluation will always be severely limited. For one thing, the number of engagements per sortie is usually limited to three. Second, the number of sorties that can be dedicated to such evaluation purposes, especially if they

are to be conducted in a somewhat controlled manner, is limited to one or two at most. The problem is analogous to baseball. How well can you estimate a player's batting average over the season from three or four times at bat during the first game? Moreover, as Lane (1986) points out in a review article of air combat measurement, extensive research conducted during World War II indicates that day-to-day reliability of such outcome measures was "low" or "near-zero." Given the fact that we cannot observe aircrew performance over a long period of time, the challenge is to find measures that are highly related to operationally relevant criteria (win/loss), but yet are not as affected by the sources of unreliability that dramatically impact readily available outcome measures. For these reasons, the present study has chosen engagement outcome as the primary criterion against which to validate measures of performance.

Review of Candidate Measures

Discussion now turns to the other side of the equation, namely, the candidate measures of performance. A fair amount of work has been done in this area over the past 15 years. Excellent reviews have been published (Briedenbach et al., 1985; Kelly, 1988; Brecke & Miller, 1991) and for that reason, no attempt will be made to review that literature in any detail. Despite the divergent purposes for which candidate measures have been developed and tested, they can be placed into four fairly distinctive categories. These include measures of positional advantage, weapons employment, aircraft state, and energy management. Each of these categories will be briefly summarized in the following sections.

Positional Advantage Measures

Positional advantage measures attempt to describe the geometric relationship of two aircraft in space in terms of their tactical advantage. Oberle (1974) began the development of such measures by proposing a discrete tactical state model that reflected positional dominance and vulnerability. Later, Simpson (1976) proposed a measure known as the "Performance Index" which attempted to continuously assess the tactical value of the interaircraft geometry and to reward maneuvering to a rule of thumb, rear hemisphere weapon envelope. Stacy (1980) introduced a measure known as the "PK-Index" which consisted of the dynamic, evolving value of the kill probability of each onboard weapon. The PK-Index assigned a weapon kill value to the evolving interaircraft geometry. Since the computation was carried out simultaneously for the adversary, a vulnerability measure was also assigned dynamically. The All-Aspect Maneuvering Index (AAMI), which is evaluated in the present study, was proposed by McGuinness, Forbes, & Rhoads (1984) as an extension of the Performance Index that is designed to reflect employment of all-aspect weapons.

Weapons Event Measures

A second category of measures concern the occurrence of discrete events while employing the weapons system. Perhaps the most well-developed example of such measures is the Performance Assessment and Appraisal System (PAAS) (Ciaverelli, Williams, & Pettigrew, 1981). Specific objectives for which measures are generated include such events as radar contact, radar lock, visual identification, first shot, first kill, parameters at weapons release, and engagement outcomes.

Aircraft State Measures

A third category of measures concern the state of the aircraft throughout the engagement. Typically, specific state parameters such as airspeed, vertical velocity, etc. are summarized throughout the entire engagement. The most significant whole engagement measure validation attempt found in the past literature was a study by Kelly, Wooldridge, Hennessy, Vreuls, Barnebey, Cotton, & Reed (1979). These authors described how well a set of candidate performance measures discriminated between expert, intermediate, and novice F-4 fighter pilots. In this study, pilots were categorized into three distinct groups, depending on past experience levels. These pilots flew against one another in a round-robin format, with six subjects matched against one another over a one-week testing period. Every pilot fought each of the other five pilots in three successive engagements. Over 400 engagements were recorded using this procedure. Univariate and multivariate analyses were then employed to discover significant measures. The result was a discriminant function model that classified pilots with over 92% accuracy into the three experience level criterion groups.

Energy Management Measures

The final category of measures are concerned with the pilot's control of the energy loss/gain characteristics of the aircraft and yield metrics such as Specific Excess Energy (P_s), available maximum and sustained turn rates and scaled metrics such as the Energy Management Index (EMI). It has been suggested that Energy management is one of the most important aspects of ACM success (Deberg, 1977).

Selection of Candidate Measures

Selection of the initial set of candidate measures to be tested against the criterion of engagement outcome was based upon both theoretical and practical concerns. Since the principal concern of the investigation was to develop measures that could be used for future transfer of training evaluations, the most important limitation is that the measures be available in-flight. For all practical purposes, this translates into the current limitations of data available from the ACMI. Moreover, to increase the reliability of the measure set, we

decided to limit measures to those that are continuously computed throughout the engagement. Both of these requirements effectively served to eliminate the discrete weapons event measures from consideration in this study. As discussed earlier, discrete events are notoriously unreliable from a psychometric standpoint, and from a practical standpoint, many are simply not available on today's ranges. For example, even the simplest of events such as "trigger squeeze" or "missile launch" are often not available for the F-15.

Selection of candidate measures was also affected by the equipment used for data gathering. As will be described in detail later, all engagements were flown in the SAAC located at Luke AFB, Arizona. The raw data were gathered and summary measures computed using the Air Combat Maneuvering Performance Measurement System (ACM PMS). For example, the AAMI was the positional advantage measure of choice since it is computed real time within the ACM PMS. A final consideration was results from previous investigations. For that reason, some of the whole engagement summary measures found by Kelly et al. (1979) to successfully discriminate pilot experience level were included within the initial analysis. Based on these considerations the final candidate set of measures is presented in Table 1. The aircraft state measures were simply the means of the parameters computed over the length of the engagement. The AAMI Offensive score was the mean AAMI computed over the length of the engagement while the AAMI Defensive was simply the AAMI Offensive score of the opponent. AAMI % > 0 was the percentage of time the AAMI score of the subject was positive while AAMI Mean > 0 was the average of those samples wherein the AAMI score was positive. A complete description of the basic AAMI calculation can be found in McGuiness, Forbes, & Rhoads (1984).

Table 1. Candidate Performance Measures

	AIRCRAFT STATE	POSITION ADVANTAGE
VERTICAL VELOCITY	X	
ROLL RATE	X	
TURN RATE	X	
G LOADING	X	
AIRSPEED	X	
CLOSING VELOCITY	X	
AAMI OFFENSIVE		X
AAMI DEFENSIVE		X
AAMI % > 0		X
AAMI MEAN > 0		X

Experimental Rationale

The questions posed in this study required the collection of an ACM engagement database under highly controlled conditions. At the outset we decided that the initial database would consist of 1V1 engagements flown in the SAAC. Advantages of data collection in the simulator were: (1) greater control over experimental conditions; and (2) access to greater information regarding actual performance. Additional considerations were: (1) the size and experience characteristics of the subject sample; (2) adversary characteristics; (3) number of engagements flown and their initial conditions; and (4) other information to be gathered as part of the data collection effort. These considerations were tempered with the availability of time and scheduling in the SAAC to arrive at a protocol for the actual data collection.

Sample Characteristics

Since the SAAC simulates both the F-15 and F-16 aircraft using interchangeable cockpits, we decided to include an equal sample of pilots who are current in each of two aircraft. An attempt was also made to include a wide range of experience levels for each aircraft. At one end of the continuum would be recent Replacement Training Unit (RTU) graduates with very limited experience at the aircraft. At the other end would be operational pilots with a large number of hours (>1000) in the aircraft. The initial sample size was based on the statistical convention of at least 5 observations per candidate measure.

Adversary Characteristics

One characteristic of ACM that is problematic for research investigations of this type is that the pilot must constantly react to a maneuvering adversary. Success, to a large degree, depends upon the competency of that adversary. Yet, for control purposes, it is necessary to establish some type of standard adversary. Computer-driven adversary models are available, but it is often questionable the extent to which they model real-world behavior. The alternative is to have a single human adversary. Aside from practical considerations, there is still the problem of generalization. For the purposes of this investigation, we decided that lack of standardization was a far greater problem than the potential for poor generalization. For that reason, we decided to use both a computer-driven adversary model and human pilot during data collection. The adversary model used was the AML as implemented in the SAAC. The human pilot, who served as the subject-matter expert (SME) for this research investigation is a former U.S. Air Force fighter pilot, Fighter Weapons School graduate, with 3,870 flight hours, the majority of which were in the F-4. They also include over 400 combat hours in Southeast Asia. To avoid confusion, the human pilot will be referred to as the "SME" in future sections.

Engagement Characteristics

We decided that all engagements would be 1V1 with opponents starting in a neutral position. Half of the engagements would be flown starting in a head-on starting position while the other half would be flown starting from a line abreast position. For the line abreast starting position, half would be initiated with adversary on the right while the other half would be initiated with adversary on the left.

Additional Data

A variety of other information was gathered as part of the data collection effort since it was envisioned that the resulting engagement database would serve a number of purposes. In many cases, the data are not pertinent to the specific objectives of this study. Included were: (1) a Pilot History Questionnaire which gathered information on previous flight experience; and (2) an ACM Tactics Knowledge Questionnaire, a 25-item paper-and-pencil multiple choice test attempting to measure knowledge of tactics. After all engagements were flown, the SME completed a questionnaire designed to evaluate the subject's flying ability along a number of dimensions considered pertinent to air combat.

Specific Research Objectives

To summarize, the goal of the present research was to produce a composite measure of continuous performance that is highly predictive of engagement outcome. Specific objectives included:

1. To determine the effects of the experimental factors, i.e., aircraft type, type of opponent, starting position, and order of presentation, on engagement outcome.

2. To determine the statistical relationship of each candidate measure to each experimental factor.

3. To determine the relative importance of each measurement category (i.e., positional advantage and aircraft state) for the prediction of engagement outcome.

4. To produce a composite measure of performance from all candidate measures that maximizes the prediction of engagement outcome.

METHOD

Subjects

The sample consisted of 125 U.S. Air Force fighter pilots from the Tactical Air Forces, including active duty, reserve, and guard units. Of this sample, 60 were current in the F-15 and 65 were current in the F-16. The average age of the sample was 31.3 years. Average total flight time was 1,867 hours with an average fighter time of 1,009 hours. Average total time in the F-15 or F-16 was 435 hours.

Equipment

The SAAC, in conjunction with the ACM PMS, was used for gathering all engagement data. The SAAC is located at Luke Air Force Base, AZ, and is used for training air-to-air combat. It has two F-15 cockpits and two F-16 cockpits to be used interchangeably in either of the two domes. Each cockpit contains a G-seat and a G-suit hook-up. The field of view presented is 296 degrees horizontal by 150 degrees vertical and is controlled by a common Gould Model 32/87 mainframe computer system which allows free flow air combat engagements between two pilots. The computer is also programmed to fly an independent opponent, the AML, so that a 1 v 1 v 1 or a 2 v 1 engagement can be conducted. The Instructor/Operator Station (IOS) is also designed to permit the instructor to fly the adversary from the console, although that capability was not used in this investigation. The IOS also gives a computer-generated graphic 3-D depiction of engagements on a Silicon Graphics Iris Model 2400 Turbo Graphics Display, as well as out-of-cockpit views from both aircraft.

The ACM PMS is interfaced with the SAAC via ethernet connection and gathers engagement data passively from the SAAC. The system is controlled by a single Gould Model 32/67 mainframe computer and uses a Silicon Graphics Iris Model 2400 Turbo Graphics Display. This system has been installed and extensively tested for the purposes of research data collection. Operation of the ACM PMS requires that one research staff member be present during engagements to direct computer inputs.

Procedure

The experiment was administered in groups of two. Subjects were first given a verbal description of the study, in which they were told that they would be flying eight ACM engagements against a computer-generated adversary. They were told that they would be fighting a computer adversary from three neutral start positions, and that they would be informed where the bogey is before the beginning of each engagement. Subjects were then escorted into the simulator for a brief explanation of the simulator specific equipment associated

with the F-15 or F-16 cockpit. Once the subjects were in the simulator they were given an orientation of simulator capabilities and handling characteristics.

As part of the orientation, subjects flew a 5-minute "free play" engagement with the AML to give them a feel for flying the simulator in a dynamic environment. At the conclusion of the orientation flight subjects were split up. Subject 1 remained in the simulator to fly the engagements while Subject 2 was escorted to the debrief room to fill out the Pilot History Questionnaire and take the Tactical Knowledge Test. Subject 1 then flew a Basic Handling Test consisting of: (1) straight-and-level flight; (2) a loop; (3) an Immelmann; and (4) a gun tracking exercise. Upon completion of the handling test, Subject 1 then flew a total of eight engagements, four against the AML and four against the human opponent. Subject 1 then egressed from the simulator and was taken to the debrief room and completed the two questionnaires. Subject 2 was brought out of the debrief room and asked to ingress into the simulator to complete the flying portion of the experiment. Subject 2 then completed the experiment using the identical procedure just described.

The following rules of engagement were in effect during the simulated engagements. Both aircraft had an unlimited supply of fuel and ordnance (air-to-air heat missiles and 20-MM gun rounds). To make the data comparable, F-15s were not permitted usage of their radar missiles. Furthermore, normal peacetime flying safety rules were not in effect so that high angle/front quarter gunshots and low altitude fighting were permitted.

Engagements continued for 3 minutes or until a kill, whichever occurred first. If a kill was the result of impact with the ground, an over-G or a midair collision, elapsed time determined the status. If the elapsed time was less than 2 minutes, the engagement was reflown using the same starting conditions. However, if the engagement went for over 2 minutes, the data was retained and the subject was allowed to proceed to the next engagement. Engagements over 2 minutes that ended in midair collisions were recorded as a draw, while impact with the ground or over-G kills were recorded as a win for the surviving aircraft. The decision to use the 2-minute cut-off was based on time limitations for use of the simulator. Since the SAAC is also used in support of operational training, it was necessary to ensure that data collection was completed within the allocated time block.

Four randomized engagement orders were prepared in an attempt to control for potential order effects resulting from the use of different adversaries and starting positions. The randomization was constrained so that all subjects flew the eighth and final engagement against the AML. This order enabled the SME to egress from his cockpit without the subject identifying who his opponent had been. The purpose of this order was to prevent the possibility of biasing future subjects who may be informed that they were flying against a highly skilled former Air Force pilot during the experiment.

Three different initial set-ups were used, which all began with both aircraft at an altitude of 15,000 feet and at an airspeed of 400 knots. The three

set-ups were with the opponent or "bogey" on the right-hand side (two engagements), bogey on the left-hand side (two engagements), and bogey head-on (four engagements). For head-on engagements only, the restriction was made that no ordnance could be employed prior to the first merge, to prevent quick head-on missile shots. All engagements began by confirming that the subject had a visual contact or "tally ho" on the bogey (side by side) or a radar lock-on (head-on starts). Once the subject responded that he did, the experimenter replied, "Ready, ready, the fight is on."

Data from the SAAC were transferred in real time and recorded onto the ACM PMS. All candidate measures of performance were calculated within the ACM PMS and transferred to floppy disc. Data were then transferred to a Digital Equipment Corporation VAX 6310 for statistical analysis.

Experimental Design

In design terms, there were five factors in this investigation. Two were between-subjects factors, Aircraft and Order. For Aircraft, there were two levels, F-15 and F-16. For Order, there were four levels corresponding to the four different orders of presentation. There were three within-subjects factors, Starting Position, Opponent, and Outcome. For Starting Position, there were three levels, Head-on, Bogey on the Right, and Bogey on the Left. For Opponent, there were two levels, the AML and the Human opponent.

The Outcome factor requires discussion since it represents both a dependent and independent variable, depending upon the particular analysis. Each engagement was categorized as either a Win, Draw, or Loss. For those analyses attempting to predict Outcome or to determine its relations to the other experimental manipulations, it is used as a dependent variable. However, for those analyses attempting to determine the effects of each of the experimental factors including the Outcome category upon the candidate measures, it is used as an independent variable. This distinction should become readily apparent in the Results section of the report.

RESULTS

Experimental Effects on Engagement Outcome

An attempt was made to determine if any of the experimental manipulations (i.e., Aircraft, Order, Starting Position, and Opponent) were significantly related to Outcome. Frequency tables were constructed crossing each factor against Outcome. Chi-square tests were performed to determine the independence of these relations. The results are presented in Table 2. As shown, only the Opponent factor was significantly related to Outcome. In fact, it has a very large effect. In essence, the two opponents represent two distinct levels of difficulty. The data clearly indicate that pilots in our sample defeat the AML

most of the time (77.8%) and conversely are defeated most of the time by our human opponent (72.4%). From these analyses, it was clear that Opponent was a very potent factor within the data and must be carefully considered in subsequent analyses. Although not achieving significance in terms of a traditional $\alpha = .05$, the Order effect did approach statistical significance. As shown in Table 2, the primary effect appeared to be a reduced frequency in the Draw category for Orders 3 and 4.

To further investigate the relationship of these experimental manipulations to Outcome, a logistic regression analysis was performed (Moran, Engleman, Fitzgerald & Lynch, 1990). The dependent variable was Outcome while the independent variables in the model consisted of Aircraft, Order, Starting Position, and Opponent as well as their first-order interactions. Using a stepwise procedure, the final model produced by the analysis consisted of only two variables, Opponent and Order. None of the first-order interactions were included within the final equation indicating no significant interaction effects among these categorical variables. Approximate chi-squares and their associated probabilities were 580.20 ($p < .0001$) and 12.82 ($p = .0459$) for the Opponent and Order effects, respectively. These results further confirm the importance of the Opponent factor and suggest a small Order effect.

Effects of Individual Candidate Measures

The next step was to examine the effects of the experimental factors on each of the candidate measures, with particular emphasis on engagement outcome. In these analyses, the primary intent was to determine the unique contribution of each factor. Main effect descriptive statistics were computed for each candidate measure and are reported in Appendix A. The contribution of each experimental factor to the candidate measures was determined by an analysis of variance (ANOVA) procedure. Since the primary concern was the impact on Outcome, the approach for constructing the analysis model was to examine the contribution of Outcome after the variance due to all other effects had been removed. In other words, the approach was to test the significance of what might be termed the residual variance for each effect. The actual analyses were conducted on a statistical package entitled RUMMAGE (Bryce, 1980) since it enabled these types of tests. A five-factor model was constructed that tested only the main effects and first-order interactions. A subject factor was also included for testing the two between-groups factors, Aircraft and Order. Additionally, pairwise comparison tests were computed using the Scheffe' technique with $\alpha = .05$. Given the large sample size, the Scheffe' was chosen since it is more conservative than either the Least Significance Difference (LSD) or Bonferonni procedures that were also available in the analysis package. An overall summary of the results are presented in Table 3 with the detailed results of the ANOVAs for the two categories of measures in Tables 4 and 5.

Table 2. Observed Frequencies for Engagement Outcomes

FACTOR	OUTCOME			
OPPONENT	**WIN**	**DRAW**	**LOSE**	**TOTAL**
AML	389	79	32	500
HUMAN	67	71	362	500
	χ_2 (df=2) = 504.2 (p<.0001)			
AIRCRAFT	**WIN**	**DRAW**	**LOSE**	**TOTAL**
F-15	224	60	196	480
F-16	232	90	198	520
	χ_2 (df=2) = 4.56 (p=.10)			
VERSION	**WIN**	**DRAW**	**LOSE**	**TOTAL**
ORDER 1	107	50	91	248
ORDER 2	115	44	97	256
ORDER 3	120	26	110	256
ORDER 4	114	30	96	240
	χ_2 (df=6) = 12.32 (p=.06)			
START POSITION	**WIN**	**DRAW**	**LOSE**	**TOTAL**
HEAD ON	221	78	202	501
BOGEY RIGHT	113	38	99	250
BOGEY LEFT	122	34	93	249
	χ_2 (df=4) = 1.67 (p=.80)			

Table 3. Summary of Significant Effects for Candidate Measures

MEASURE	EFFECT														
	A	V	S	O	R	AV	SO	VO	OR	VS	AS	SR	AO	VR	AR
VERTICAL VELOCITY				***				*	**						
ROLL RATE			***												
TURN RATE	**		***	***	***							*			*
G LOADING	**			***						*			**		
AIRSPEED	**		***		***							*			
CLOSING VELOCITY			***	***	***		***		***			***			
AAMI OFFENSIVE		*	***	***	***				**						
AAMI DEFENSIVE			***	***	***		***					***			
AAMI % > 0		*	***	***	***								*		
AAMI MEAN 0			***	***	***		*	*	***		*	*			

*p < .05
**p < .01
***p < .001
A-AIRCRAFT
V-ORDER
S-START POSITION
O-OPPONENT
R-OUTCOME

Order Effects

As shown in Tables 3, 4, and 5, significant effects occurred for only two measures, both of which were in positional advantage category. However, none of the pairwise comparison tests reached significance.

Aircraft Effects

As shown in Tables 3, 4, and 5 a number of aircraft state measures were significantly affected by aircraft type. However, none of the positional advantage measures showed any significant effects. Table 6 presents the means for those measures producing a statistically significant effect. As indicated the F-16 demonstrated higher turn rates, while the F-15 produced higher Gs and airspeeds.

Start Position Effects

A large number of significant effects were observed among the candidate measures as a result of Start Position. In fact, virtually all measures showed significant effects across the three categories. Table 7 presents the means for these measures, as well as which pairwise comparisons were significant.

Table 4. F-Ratios for Aircraft State Measures

EFFECT	DF	VERTICAL VELOCITY	ROLL RATE	TURN RATE	G LOADING	INDICATED AIRSPEED	CLOSING VELOCITY
BETWEEN-GROUP FACTORS							
AIRCRAFT(A)	(1,117)	0.08	0.60	7.35**	10.20**	10.02**	0.28
ORDER(V)	(3,117)	0.47	1.56	0.45	0.89	0.57	0.45
AxV	(3,117)	0.18	0.60	0.83	0.79	1.04	2.47
WITHIN-GROUP FACTORS							
START (S)	(2,842)	2.45	11.96***	27.65***	1.77	9.22***	106.21***
OPPONENT (O)	(1,842)	43.15***	1.19	199.56***	22.92***	2.66	51.29***
RESULT (R)	(2,842)	0.35	2.27	7.57***	2.57	7.01***	22.21***
SxO	(2,842)	0.98	1.59	0.26	0.58	0.10	13.06***
VxO	(3,842)	2.66*	0.51	0.88	0.29	0.50	1.65
OxR	(2,842)	4.72**	0.03	1.28	1.88	2.62	27.18***
VxS	(6,842)	0.58	0.30	0.85	2.19*	1.13	1.8
AxS	(2,842)	0.89	1.21	2.14	0.92	0.06	0.61
SxR	(4,842)	0.49	1.99	2.89*	1.52	2.64*	6.13***
AxO	(1,842)	1.27	1.03	0.70	8.25**	1.60	3.44
VxR	(6,842)	1.43	0.18	1.55	1.70	0.65	1.73
AxR	(2,842)	0.25	0.96	4.04*	2.60	0.10	0.39

*$p < .05$

**$p < .01$

***$p < .001$

Table 5. F-Ratios for Positional Advantage Measures

EFFECT	DF	AAMI OFFENSIVE	AAMI DEFENSIVE	AAMI % > 0	AAMI MEAN > 0
BETWEEN-GROUP FACTORS					
AIRCRAFT(A)	(1,117)	0.31	1.92	0.10	0.94
ORDER(V)	(3,117)	2.83*	2.18	3.15*	1.96
AxV	(3,117)	0.56	0.93	1.18	0.20
WITHIN-GROUP FACTORS					
START (S)	(2,842)	16.36***	13.11***	5.07**	46.21***
OPPONENT (O)	(1,842)	66.35***	662.23***	22.96***	96.54***
RESULT (R)	(2,842)	97.65***	68.47***	64.57***	90.26***
SxO	(2,842)	2.11	11.27***	1.30	4.53*
VxO	(3,842)	1.35	1.88	0.56	2.68*
OxR	(2,842)	5.96**	0.86	0.73	8.50***
VxS	(6,842)	0.65	0.46	0.43	1.69
AxS	(2,842)	2.53	0.92	2.22	3.28*
SxR	(4,842)	0.73	9.75***	1.20	2.53*
AxO	(1,842)	2.86	3.00	3.87*	0.86
VxR	(6,842)	0.52	1.74	0.55	0.94
AxR	(2,842)	1.31	1.74	2.34	0.85

*p < .05
**p < .01
***p < .001

Table 6. Means for Significant Aircraft Effects

	F-15	F-16
TURN RATE	0.22	0.23
G LOADING	3.71	3.37
AIRSPEED	306.70	278.23

Table 7. Means and Significant Pairwise Comparisons for Starting Position

	1 HEAD_ON	2 BOGEY_R	3 BOGEY_L	1/2	1/3	2/3
ROLL RATE	-0.01	0.01	-0.03	X	X	X
TURN RATE	0.22	0.24	0.23	X	X	X
AIRSPEED	305.14	279.87	277.32	X	X	
CLOSING VELOCITY	113.61	27.04	24.92	X	X	
AAMI OFFENSIVE	36.64	31.50	31.08	X	X	
AAMI DEFENSIVE	30.93	26.49	26.46	X	X	
AAMI % > 0	56.43	54.33	53.56		X	
AAMI MEAN > 0	61.74	52.51	52.58	X	X	

Opponent Effects

Tables 3, 4, and 5 also indicate a large number of significant effects due to type of opponent for both categories of measures. Table 8 presents the means for those measures that were significant.

Table 8. Means for Significant Opponent Effects

	AML	HUMAN
VERTICAL VELOCITY	37.63	95.46
TURN RATE	0.20	0.25
G LOADING	3.37	3.70
CLOSING VELOCITY	31.59	108.18
AAMI OFFENSIVE	45.57	22.37
AAMI DEFENSIVE	12.28	45.13
AAMI % > 0	64.82	45.56
AAMI MEAN > 0	68.07	46.24

As shown, aircrews flying against the human opponent tended to fight more in the vertical dimension, turned harder, pulled more G's and produced higher closing velocities. For positional advantage, it is clear that higher offensive scores were achieved against the AML and higher defensive scores were produced when flying against the human opponent. Again, this tends to corroborate the importance of the Opponent factor within this engagement database.

Outcome Effects

A number of significant effects also occurred for the main factor of interest, Outcome. Significant differences were found for measures from both categories. Means and significant pairwise comparisons are presented in Table 9. For state measures, the results indicated Wins to have significantly lower turn rates and higher airspeeds while Losses have significantly higher closing velocities. For the positional advantage measures, which are also presented in Figure 1, there were reliable differences for all measures across all outcomes, with the exception of the AAMI Defensive score which only discriminates Losses from the other two outcomes.

Table 9. Means and Significant Pairwise Comparisons for Outcome

	1 WIN	2 DRAW	3 LOSE	1/2	1/3	2/3
TURN RATE	0.21	0.23	0.25	X	X	
AIRSPEED	303.80	273.31	285.19	X	X	
CLOSING VELOCITY	47.32	28.99	111.57	X		X
AAMI OFFENSIVE	47.60	28.16	20.41	X	X	X
AAMI DEFENSIVE	14.78	23.96	46.63		X	X
AAMI % > 0	67.20	51.58	42.67	X	X	X
AAMI MEAN > 0	69.43	51.10	45.24	X	X	X

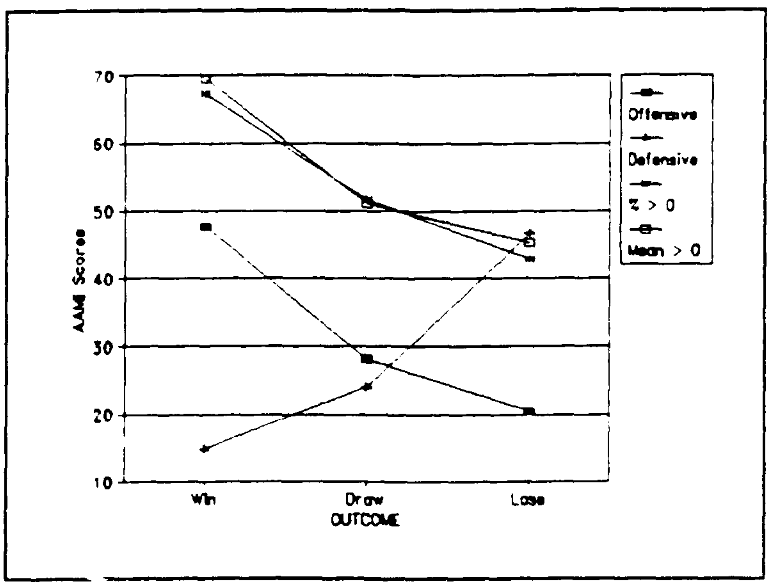

Figure 1. Measures of Position Advantage as a Function of Engagement Outcome.

First-Order Interactions

As shown in Tables 3, 4, and 5, there were a number of significant first-order interactions. Of the ten first-order interactions, two produced no significant effects. These included Order by Aircraft and Order by Outcome. The Order by Start Position interaction provided an overall effect for G. However, none of the pairwise comparisons were significant. The remaining significant interaction effects are presented in Figures B-1 through B-18 of Appendix B. The Order by Opponent interaction, produced two significant effects, Vertical Velocity and AAMI Mean > 0. For vertical velocity (see Figure B-1), pairwise comparison tests indicated significant differences between the AML and Human opponent, for Orders 1 and 2 but not Orders 3 and 4. For the AAMI Mean > 0 measure (see Figure B-2), significant differences were found for all orders except Order 3.

For Aircraft by Start Position, one measure produced a significant effect, AAMI Mean > 0. It is plotted in Figure B-3 of Appendix B. Pairwise tests indicated no significant aircraft effects across the three starting positions. The only significant tests were between starting positions (head-on vs. line-abreast starting positions) for both aircraft, which are consistent with the main effects results.

For Aircraft by Opponent, two measures produced significant effects, G and AAMI % > 0. For G (see Figure B-4), pairwise tests indicated significantly higher G's by the F-15 against the human opponent. For the F-16, there were no differences due to the type of Opponent. However, there were no differences against the AML. For AAMI % > 0 (see Figure B-5), significant differences occurred for the F-15, but not F-16 due to type of opponent. For this particular AAMI measure, the data indicate significantly lower scores against the Human opponent for the F-15.

For Aircraft by Outcome, only Turn Rate produced a significant effect. It is plotted in Figure B-6 of Appendix B. Significant aircraft differences occurred for Losses, but not for Wins or Draws. In that case, the F-15 had a lower turn rate for losses. For the F-16, significant differences occurred between Wins and Losses. However, none occurred for the F-15.

For Starting Position by Opponent, three measures produced significant effects, Closing Velocity, AAMI Defensive, and AAMI Mean > 0. For V_c (see Figure B-7), a significant difference occurred between Opponent only for the head-on start position. For AAMI Defensive (see Figure B-8), a significant difference occurred between the Head-on and Bogey-Right positions only for the Human opponent. For AAMI Mean > 0 (see Figure B-9), significant differences occurred between the Head-on and both Bogey-Left and Bogey-Right start positions against the Human opponent. Against the AML opponent, significant differences occurred only between the Head-on and Bogey-Right start positions. For all start positions, there were significant differences between the two Opponents.

For Start Position by Outcome, five measures produced significant effects, Turn Rate, Airspeed, V_c, AAMI Defensive, and AAMI Mean > 0. For Turn Rate (see Figure B-10), significant differences occurred between the head-on and the two line-abreast starting positions only for losses. No differences occurred for either Wins or Draws. For Airspeed (see Figure B-11), a significant difference occurred only between the head-on and one of the line-abreast starting positions only for losses. A significant difference occurred between wins and draws for the head-on start position. For V_c (see Figure B-12), significant differences occurred between the Win/Draw and Lose/Draw comparisons only for the head-on starting position. Significant Head-on vs. line-abreast differences also occurred for Wins and Losses, but not for Draws.

For AAMI Defensive (see Figure B-13), significant differences between the Head-on and Bogey-Right/Bogey-Left start positions were obtained for Wins but not for Draws or Losses. For the Bogey-Right start position, significant differences were obtained between Wins and Losses, but not between Draws and Losses. All other differences were consistent with the main effect findings. For AAMI Mean > 0 (see Figure B-14), significant Win/Draw and Win/Lose effects were obtained for all starting positions. However, none of the Draw/Lose comparisons were significant. Significant Head-On/Bogey-Right and Head-On/Bogey-Left comparisons were obtained for Wins and Losses, but not for Draws. No significant Bogey-Right/Bogey-Left comparisons were obtained; this finding is consistent with the main effect findings.

And finally, for the Opponent by Outcome interaction, four measures produced significant effects: Vertical Velocity, V_c, AAMI Offensive and AAMI Mean > 0. These are plotted in Figures B-15 through B-18 of Appendix B. For Vertical Velocity (see Figure B-15), significant differences occurred between the AML and Human opponent only for Losses. For V_c (see Figure B-16), significant differences occurred between the AML and Human opponent only for Wins. Likewise, significant Win/Draw, Win/Loss, and Draw/Loss comparisons occurred for the Human opponent, but no differences occurred when flown against the AML. For AAMI Offensive (see Figure B-17), all comparisons were significant except the AML/Human contrast for losses and the Draw/Loss contrast for the Human opponent. The same pattern of results was also found for the AAMI Mean > 0 measure (see Figure B-18).

Development of Composite Measures

The initial set of analyses focused on determining the effects of the experimental manipulations upon the individual candidate measures. Those results also helped shape the types of analyses performed in addressing the main concern of the study, namely the development of a composite measure of performance that is predictive of engagement outcome. Perhaps the most salient result from analysis of the candidate measures was the very powerful effect of Opponent type. For that reason, all multivariate analyses were conducted separately for each type of Opponent. The approach taken was to first determine the relation of each measurement category (i.e., aircraft state,

and positional advantage) to engagement outcome and then to develop a composite measure using only those categories found to be significantly predictive of engagement outcome. As previously discussed, outcomes were coded as either a Win, Draw, or Loss. In addition to using these categories, additional analyses were performed using the criterion of simply Win or Loss. Since the outcome of Draw occurred in only 15% of the entire sample of 1,000 engagements, the net effect on sample size was fairly small. The statistical technique used in these analyses was discriminant analysis. The specific software package used was the BMDP Stepwise Discriminant Analysis (Jennrich & Sampson, 1990).

Aircraft State Measures

The first category included measures of aircraft state including Vertical Velocity, Turn Rate, Airspeed, G, and V_c. For engagements flown against the AML using the Win/Draw/Lose criterion, the discriminant analysis produced a three-variable equation (Turn Rate, Airspeed, V_c). Using this equation, only 42% of the engagements were correctly classified, which is only a few percentage points greater than chance. For engagements flown against the Human opponent, a two-variable equation (Vertical Velocity, V_c) was produced which correctly classified 49% of the engagements. Interestingly enough, the equation correctly classified 94% of the Draws, but only 36% of the Wins and 43% of the Losses.

For the Win/Lose criterion, a four-variable equation (Turn Rate, V_c, Airspeed, Vertical Velocity) was produced for engagements flown against the AML that correctly classified 87% of the engagements. Percentages correctly classified in each category were roughly the same with 89% for Wins and 85% for Losses. For engagements flown against the Human opponent, only one variable entered the equation (Vertical Velocity) which produced 66% correct classifications. Using this equation, 69% of the Losses were correctly classified, whereas only 43% of the Wins were correctly classified which is below chance level.

Positional Advantage Measures

The second category included measures of positional advantage as estimated by the four AAMI scores. For engagements flown against the AML using the Win/Draw/Lose criterion, all four measures entered the equation and correctly classified 72% of the outcomes. Percent correct classifications included 79% of the Wins, 75% of the Losses, but only 38% of the Draws. For engagements flown against the Human opponent, again all four variables entered the equation leading to 64% of the outcomes being correctly classified. Percent correct classifications included 73% of the Wins, 63% of the Draws, and 63% of the Losses.

For the Win/Lose criterion, again all four variables entered the equation for engagements flown against the AML, leading to 97% correct classification of all outcomes. These included 98% of the wins and 96% of the losses. For

engagements flown against the Human opponent, again all variables entered the prediction equation resulting in an overall 83% correct classification. These included 81% of the Wins and 84% of the losses. Of the two categories of measures, it is readily apparent that positional advantage is the most important and predictive of engagement outcome.

Aircraft State and Positional Advantage Measures

For the final set of analyses, candidate measures from the Aircraft State and Positional Advantage categories were permitted entry into the discriminant analysis. For the Win/Lose/Draw criterion, a five-variable equation (AAMI Defensive, AAMI % > 0, AAMI Mean > 0, Turn Rate, G) was produced for engagements flown against the AML which correctly classified 71% of the outcome. These included 74% of the Wins, 75% of the Losses, but only 52% of the Draws. For engagements flown against the Human opponent, a six-variable equation (all AAMI measures plus Vertical Velocity and V_c) was produced which correctly classified 70% of the outcome. These included 78% of the Wins, 78% of the Draws, and 68% of the Losses.

For the Win/Lose criterion against the AML, a six-variable equation (all AAMI measures plus V_c and Airspeed) was produced which correctly classified 97% of the outcomes. These included 99% of the Wins and 96% of the Losses. For engagements flown against the Human opponent, a six-variable equation *(all AAMI measures plus Turn Rate and V_c)* was produced which correctly classified 84% of the outcomes. These included 78% of the Wins and 85% of the Losses.

An overall summary of these analyses is presented in Table 10. From these data, several things are readily apparent. In terms of the two criteria used for engagement outcome, it is clear that predictions of group membership are much higher for Win vs. Lose. Moreover, predictions for outcome are generally higher for those engagements flown against the AML, especially when the criterion is Win/Lose.

Comparison of AML and Human Models

Because of the large differences as a function of Opponent, we decided to conduct separate analyses for predicting outcomes against the AML and Human. However, the issue of the robustness of the resulting composite measures still remains. Of most interest is the question of whether measures predictive of outcome against the AML are also predictive of outcome against the Human. To answer this question, the classification function developed for engagements flown against the AML was applied to engagements flown against the Human opponent and vice versa. For the AML function applied to Human engagements, overall 73.2% of the engagements were correctly classified, which is actually a little better than the model applied to the original data. However, the percent correct classifications for each category were 4.5% for Wins, 2.9%

for Draws, and 99.7% for Losses. For the Human classification function applied to AML engagements, similar results were obtained. Overall, 72.0% of the outcomes against the Human opponent were correctly classified. However, the percent classifications for each category were 85.3% for Wins, 34.2% for Draws, and 3.2% for Losses. From these findings, it was clear that one cannot adequately classify outcomes for Human engagements based on classification functions derived from AML engagements and vice versa.

Table 10. Percent Correct Classifications for All Functions

	WIN/DRAW/LOSE OUTCOME				WIN/LOSE OUTCOME		
	TOTAL	WIN	DRAW	LOSE	TOTAL	WIN	LOSE
AML OPPONENT							
AIRCRAFT STATE	42.4%	40.4%	54.4%	37.5%	87.2%	89.2%	85.3%
POSITION ADVANTAGE	72.2%	78.9%	38.0%	75.0%	96.9%	98.2%	95.7%
COMPOSITE	70.8%	74.3%	51.9%	75.0%	97.1%	98.5%	95.7%
HUMAN OPPONENT							
AIRCRAFT STATE	49.4%	35.8%	94.4%	43.1%	65.5%	43.3%	69.3%
POSITION ADVANTAGE	65.0%	73.1%	66.2%	63.3%	83.3%	80.6%	83.8%
COMPOSITE	70.4%	77.6%	77.5%	67.7%	83.5%	77.6%	84.5%

However, part of this difference may have been due to the fact that the variables included within the two models were somewhat different. Since the Position Advantage measures as a category were most predictive of engagement outcome, similar analyses were conducted in which the four AAMI measures were forced into the classification function. Analyses were conducted for both the Win/Draw/Lose criterion and the Win/Lose criterion. Similar findings emerged as seen in Table 11.

As shown, the ability of each function to predict the outcome of engagements against the opposite opponent is limited. It is of interest to note that in all instances a reasonably good classification model can be developed for each data set. In other words, the four AAMI positional advantage measures do a reasonable job of predicting engagement outcome for each type of opponent. However, when an attempt is made to predict outcome for the engagements flown against the other opponent, the accuracy of classification declines dramatically, especially for individual categories. For example, when AML data is used to predict outcomes against the human opponent, it tends to classify almost all engagements as Losses. Conversely, when Human data is used to predict outcomes against the AML, it tends to classify almost all engagements as Wins. Again, these results corroborate the importance of the Opponent variable.

Table 11. Classification Functions Applied to Other Opponent

WIN/DRAW/LOSE CRITERION	AML TO HUMAN		HUMAN TO AML	
OVERALL	72.2%	73.4%	65.0%	73.8%
WINS	78.9%	7.5%	73.1%	88.9%
DRAWS	38.0%	0.0%	66.2%	29.1%
LOSSES	75.0%	100.0%	63.3%	0.0%
WIN/LOSE CRITERION				
OVERALL	90.5%	74.0%	83.0%	79.0%
WINS	91.8%	11.9%	85.1%	100.0%
LOSSES	81.3%	100.0%	82.6%	18.8%

In a final set of analyses, classification functions were developed for the entire sample, including engagements flown against both the AML and Human opponent. Table 12 presents the accuracy of the overall classification function as well as accuracies for each opponent. As shown, while the accuracy for the entire sample is reasonably high, there is a significant decrement when one considers either of the two subsamples. Again, the same pattern emerges as shown in previous analyses. For AML data the function predicts Wins reasonably well, but not Losses or Draws. Conversely, for Human data the function predicts Losses reasonably well, but not Wins or Draws. Again, these results further confirm the importance of the Opponent factor and the inability to develop a composite measure for the entire sample that is predictive of outcome for either of the two subsamples.

Expanded Classification Functions

The analyses just mentioned clearly indicate the inability to develop a classification function for one type of opponent that would predict outcomes for the other type. Moreover, one cannot develop a classification function from the entire data sample with reasonable prediction for either of the two subsamples. The alternative is to consider the development of classification functions wherein the outcomes against the AML and Human opponent are considered as separate groups or categories. In other words, for the Win/Draw/Lose criterion, there would be six groups, three for the AML data and three for the Human data. The results of such an analysis for the entire set of measures are shown in Table 13. Candidate measures included within the model were Vertical Velocity, G, Airspeed, V_c, AAMI Offensive, AAMI Defensive, and AAMI % > 0. A similar analysis was also computed for the Win/Lose criterion. These results are shown in Table 13. Candidate measures included within this analysis were Turn Rate, V_c, and the four Position Advantage measures.

Table 12. Classification Accuracy for Functions Based on Entire Sample

WIN/DRAW/LOSE CRITERION	OVERALL	AML	HUMAN
OVERALL	73.2%	70.4%	76.0%
WINS	70.4%	80.5%	11.9%
DRAWS	45.3%	44.3%	46.5%
LOSSES	87.1%	2.5%	93.6%
WIN/LOSE CRITERION			
OVERALL	91.1%	94.8%	87.4%
WINS	87.5%	99.0%	20.9%
LOSSES	95.2%	43.8%	99.7%

As shown, the classification accuracy of these functions is less than the accuracy of those models developed specifically for each of the two Opponents (see Table 10). For the Win/Draw/Lose criterion model, 65.2% of all outcomes were correctly classified compared with 79.9% for the Win/Lose criterion model. Despite the overall reduction in classification accuracy, the functions do a reasonably good job of separating AML from Human engagements. In fact, of the 500 engagements flown against the human opponent using the Win/Draw/Lose criterion, only 13 or 2.6% were classified into an AML group. Likewise for AML engagements only 12 or 2.4% were classified into a Human group. As shown in Table 14, a similar pattern emerges for the Win/Lose criterion. Again, these results further emphasize importance of the Opponent factor and the fact that these data do represent two independent distributions with very little overlap.

Table 13. Accuracy of Expanded Six-Group Classification Function

			PREDICTED GROUPS					
			HUMAN			AML		
ACTUAL	GROUPS	% CORRECT	WIN	DRAW	LOSE	WIN	DRAW	LOSE
	WIN	67.2	45	16	2	2	1	1
HUMAN	DRAW	73.2	5	52	11	0	0	3
	LOSE	63.8	50	75	231	0	0	6
	WIN	67.1	4	1	0	261	97	26
AML	DRAW	53.2	0	0	2	22	42	13
	LOSE	65.6	1	4	0	1	5	21

Table 14. Accuracy of Expanded Four-Group Classification Function

			PREDICTED GROUPS			
			HUMAN		AML	
ACTUAL	GROUPS	% CORRECT	WIN	LOSE	WIN	LOSE
HUMAN	WIN	77.6	52	8	4	3
	LOSE	76.5	75	277	0	10
AML	WIN	83.5	6	0	325	58
	LOSE	71.9	5	0	4	23

DISCUSSION

Experimental Effects on Engagement Outcome

The first set of analyses attempted to determine whether any of the experimental manipulations significantly affected engagement Outcome. No main effects were found for either Starting Position or Aircraft. Moreover, no first-order interaction effects were found across all of the factors. However, the results clearly indicated that the Opponent factor had a major effect. As shown in Table 2, there appeared to be almost a dichotomy of results in terms of the Win/Lose categories. In general, pilots won against the computer-driven adversary model, the AML, and in turn lost against the human adversary. In other words, the AML might be considered too easy an adversary while the human might be considered too difficult. In a sense, the situation is roughly analogous to problems in test construction in terms of desired levels of item difficulty (Nunnaly, 1967). While the overall difficulty level, i.e., the probability of achieving a kill, of the two adversaries averaged together is .46 which is near the optimal of .50, the individual probabilities are .13 and .78 for the Human and AML opponents, respectively, which would not be considered within the optimal range of .40 to .60. However, it must be emphasized that the current study represents only a loose analogy to item difficulty for test construction, but nonetheless points to potential problems, especially in attempting to combine data from these two adversaries. At a minimum, it points to the necessity of either including the Opponent factor in any analyses of this data set or considering these two subsamples independently.

The results also produced an Order effect that approached statistical significance. The data indicated fewer Draws and more Losses for Orders 3 and 4. In other words, the decisiveness of these engagements appeared to be somewhat higher for some unknown reason. A comparison of the actual engagement orders provided no insight in terms of possible reasons for these observed differences.

While statistical significance is the primary criterion for judging the reliability of an effect, it gives no information as to its size (Cohen, 1965). Given the large number of observations, it is quite possible to achieve a *small*, but statistically significant effect which may be of little practical concern. To test this possibility, coefficients of contingency, designated "C," were computed for both the Opponent and Order effects (Guilford, 1965). The results were Cs of .58 and .11 for these two effects respectively. In explanation of variance terms, this translates into 33.5% and 1.2%, respectively. In other words, the Order factor, although approaching statistical significance, accounts for only a little over 1% of the variability of engagement outcome. The conclusion to be drawn is that Order had no important or practical effect on Outcome.

Effects on Individual Candidate Measures

The second set of analyses attempted to determine if any of the experimental factors produced effects upon the individual candidate measures of performance.

Order Effects

The results produced only two significant main effects due to Order, AAMI Offensive, and AAMI Mean > 0. For each of these effects, however, none of the pairwise comparison tests among the different orders reached significance. These results suggest that the four orders of presentation used during data collection produced no meaningful effects on performance.

Aircraft Effects

For Aircraft type, significant main effects were obtained for three of the state measures, Turn Rate, G, and Airspeed, but none of the positional advantage measures. These differences can best be explained through a simple comparison of the basic difference in the design of the two aircraft. The F-16 maintains a "Fly-by-Wire" or computer-controlled flight control system while the F-15 has a more conventional direct pilot input, computer-augmented flight control system. The effect of this difference is that an F-16 pilot can make maximum stick deflection inputs knowing that the computer will only move the aircraft-control surfaces an amount which will yield maximum performance for the current conditions without stalling or over G-ing the airplane. As a result, to achieve maximum performance, which is almost always desired, the F-16 pilot tends to place the throttle into full afterburner and leave it there for the duration of the engagement, relying on the automatic, computerized "limiter" to prevent a stall or over-G.

The F-15 pilot on the other hand, commands direct access to the control surfaces and can get maximum deflection on demand regardless of aircraft state. Therefore, the F-15 pilot has to be more concerned with aircraft conditions such as airspeed and G loading, since the only factor between safe maneuvering

and an over-G is his ability to properly judge current aircraft state and act accordingly. The F-15 pilot, therefore, must control his airspeed more through throttle modulation and greater use of a lower than Full A/B power setting.

There also appears to be a fundamental difference in the way F-15 and F-16 pilots approach air-to-air combat. This may be due to the aircraft differences discussed above or it may be a function of exposure to the actual air-to-air task itself. Although there is currently no objective data to support this notion, a common feeling among fighter pilots is that in addition to a cognitive understanding of the task, air-to-air combat requires a certain amount of finesse to perform well. It appears that in general the F-15 pilots tested exhibited more finesse than their F-16 counterparts. Evidence of this can be seen in that F-16 pilots tended to fly more at or near the edges of the flight envelope, relying on the F-16's excellent instantaneous performance. On the other hand, the F-15 pilots tended to stay further away from the edges of the envelope and closer to the aircraft's sustained performance parameters. For example, in both airspeed and G, the F-16 group maintained a higher maximum and the lower minimum while the F-15 group had the higher mean. This finding confirms that the F-15 pilots stayed away from the extremes, but were more able to maintain a constant and higher airspeed and G. It is purely conjecture to state, however, whether these observed differences in approach were the result of a conscious plan or were simply a function of aircraft differences.

The higher average Turn Rate of the F-16 group may also be explained in terms of the same factors just discussed. Aircraft design allows the F-16 pilot to take advantage of the computer to provide maximum turning performance for any set of conditions, so he tends to maintain full back stick pressure when turning. The result is that the computer always provides the maximum G (which results in higher Turn Rate) for the current airspeed condition. Additionally, less finesse results in the F-16 pilot pressing further into a given situation before recognizing the proper action, therefore requiring him to rely more often on maximum turn rates to recover. This results in higher peaks in turn rate throughout an engagement and subsequently a higher average Turn Rate for the entire engagement.

Start Position

Eight of the 10 candidate measures were significantly affected by starting position. As shown in Table 7, the general finding is a difference between the head-on and two line-abreast positions, but no difference between Bogey-Left and Bogey-Right positions. These general findings would certainly be expected and tend to provide an additional amount of face validity for the measures. Higher Airspeeds and V_cs would be expected from the head-on start positions since it took about 8 seconds from the start of the engagement to the actual merge. Likewise, higher Roll Rates and Turn Rates would be expected in the line-abreast start positions since the initial move for these engagements is to make a hard turn in the direction of the adversary. For Roll Rate, the significant

difference between the Bogey-Right and Bogey-Left start positions was due to the fact that the actual Roll Rate data from the SAAC is signed in terms of its direction. It seems likely that this difference would have disappeared if the absolute value of roll rate had been computed prior to computing the mean. A significant difference also occurred in Turn Rate between the line-abreast positions with higher turn rates when the bogey was on the right. The nature of such differences, however, is unknown as there appeared to be no reasonable explanation of such findings.

All positional advantage measures were influenced by Start Position with significantly higher means for the Head-on position. No differences occurred between the two line-abreast conditions. Again such findings would be expected since both adversaries will be offensive during the initial head-on pass because of the all-aspect capabilities of the AIM-9L missile. It should also be recalled that the Rules of Engagement did not permit weapon employment on the initial pass so that a relatively "fixed" amount of "offensiveness" would be added to each adversary's AAMI scores on these head-on passes.

Taken as a whole, these results indicate that Start Position does have a significant impact on these individual candidate measures. Moreover, the difference is primarily in terms of head-on vs. line-abreast with little impact of whether the Bogey is to the right or left for line-abreast start positions. For future efforts that might use this database, it seems that coding engagements in terms of head-on or line-abreast would be sufficient to account for the variability resulting from differences in Start Position. It may also be possible to eliminate such differences altogether by eliminating the data prior to the merge for head-on passes. Further work is required to determine whether such a scheme would be effective.

Opponent Effects

All position advantage measures were significantly affected by Opponent as well as 4 of the 6 aircraft state measures. The effects on the position advantage measures would certainly be expected since previous analyses had shown that Opponent had a very large impact upon engagement outcome. As shown in Table 8, the three measures of "offensiveness" were significantly higher for engagements flown against the AML. Likewise, the one measure of "defensiveness" was significantly higher for those engagements flown against the human opponent.

The observed Opponent differences for the aircraft state measures are due largely to what might be termed the "style" of the AML. As shown in Table 8, Vertical Velocity, Turn Rate, G, and V_c were all lower for engagements flown against the AML. Such differences are likely due to certain characteristics of the way the AML fights an engagement. For one thing, the AML tends to fly faster when compared with human pilots. Whenever it loses energy to a certain point, it will extend in order to increase airspeed and then reenter the fight. These characteristics are reflected in the results in that the higher

airspeeds and extended fight will tend to reduce both the Gs and Turn Rate. Because of the propensity of the AML to extend, the subject pilot will tend to "chase" his adversary thus reducing the amount of high closing velocity during an engagement. The AML also tends to fight in the horizontal rather than the vertical dimension. Again this is reflected in the results with significantly lower Vertical Velocity against the AML.

The major finding from the first set of analyses addressing the impact of the various experimental manipulations upon engagement outcome was the salience of the Opponent factor. The results of these analyses dealing with the effects of these factors upon the individual candidate measures further confirm the importance of Opponent. The findings also provide some support for the validity of the AAMI measures of positional advantage. If pilots generally "win" against the AML and "lose" against the human opponent, one could reasonably hypothesize that their "offensiveness" scores would be higher against the AML and likewise, their "defensiveness" score would be higher against the human opponent. The data from these analyses lend support to such an hypothesis, and thus provide some support for the validity of the AAMI scores. The second conclusion to be drawn from these results is that, not only are the outcomes different against the two opponents, but also that pilots fly *differently*. The characteristic differences in the way the AML fights an engagement are reflected in the aircraft state measures. These findings further reinforce the importance of this factor and the requirement that it carefully be considered in further explorations of this data set.

Outcome Effects

Discussion now turns to the major concern of the investigation, engagement outcome. The question to be answered is whether any of the individual candidate measures are systematically related to Outcome. In this instance, Outcome is being used as an independent factor only in a statistical sense. The results indicated that all of the positional advantage measures and half of the aircraft state measures showed significant differences. The three AAMI measures of "offensiveness" produced a fairly linear relationship with outcome as shown in Figure 1. The AAMI Defensive score produced significant differences between Wins and Losses, Draws and Losses, but not Wins and Draws. These findings suggest that the AAMI Defensive score, which is simply the AAMI Offensive score of the adversary, may be the best predictor of losses. Again, these findings provide further evidence for the validity of the AAMI scores.

For both Airspeed and Turn Rate, the results produced significant differences between Wins and Draws, Wins and Losses, but not Draws and Losses. In other words, these measures discriminate Wins from Draws and Losses. The measures, per se, are related in that higher airspeeds are associated with lower turn rates. It seems reasonable to speculate that Winners are better able to manage their energy and thus fly at higher airspeeds approaching the corner velocity of the aircraft in which turning is optimal. V_c produced significant differences for Wins vs. Draws, Draws vs. Losses, but not Wins vs. Losses.

In other words, average V_c discriminates Draws from Wins and Losses. These findings seem best explained by the major two types of Draws observed in this investigation. In one case, the two aircraft began a circular downward spiral which lasted until the 3-minute maximum time limit had been exceeded. In such cases, V_c during the spiral portion of the engagement approached zero. In the second case, the opponent, although very defensive, was able to jink well, thus never permitting a lethal gun tracking solution. Again, in such cases, the V_c was very small.

First-Order Interactions

Order. Of the first-order interactions involving Order, two, Aircraft by Order and Order by Outcome, produced no significant effects. Order by Start Position produced one significant effect, G, but none of the pairwise comparison tests reached significance. Order by Opponent produced significant effects for Vertical Velocity and AAMI Mean > 0. However, a perusal of the these effects, that are graphed in Figures B-1 and B-2 of Appendix B, by the SME for this investigation provided little insight as to the nature of these differences. In summary, the effects of Order were found to be few in number, inconsistent, small in size, and uninterpretable. Based on these findings, we concluded that Order had little effect on the individual candidate measures and can be disregarded in future analyses of this data set.

Aircraft by Opponent. A significant Aircraft by Opponent was found for G which is shown in Figure B-4. These results seem to be due to the combination of factors previously discussed, namely, the "style" of the AML and the "fly-by-wire" characteristics of the F-16. Since the AML flies at higher airspeeds and often extends, the Gs of its opponent are likely to be less. However, these differences are likely to be much less in the F-16 because of its flight control system that enables the pilot to achieve a maximum G for a given airspeed without getting into an over-G situation.

Aircraft by Outcome. A significant Aircraft by Outcome interaction occurred for Turn Rate which is shown in Table B-6. Although speculative, it seems reasonable that such findings are again a result of the F-16 flight control system. While losses are usually characterized by a lower energy state in general, the F-16 is still able to achieve maximum turning performance through its fly-by-wire system. For this reason, the F-16 is able to achieve a higher turn rate, even for those engagements that resulted in a Loss.

Start Position by Opponent. Three measures produced significant Start Position by Opponent interactions, V_c, AAMI Defensive, and AAMI Mean > 0. For V_c (see Figure B-7), these results are most likely due to the "style" of the AML in terms of overall higher airspeeds and the tendency to extend which leads to reduced V_cs throughout the engagement. For AAMI Defensive (see Figure B-8) and AAMI Mean > 0 (see Figure B-9), however, there appeared to be no plausible explanation of the results.

Start Position by Outcome. Three aircraft state measures produced significant Start Position by Outcome interactions, Turn Rate, Airspeed, and V_c. For Turn Rate (see Figure B-10) and Airspeed (see Figure B-11), the results are similar in that significant differences occurred across start positions only for Losses. Since Airspeed and Turn Rate are highly related, such similarities are to be expected. It seems reasonable to speculate that for head-on engagements, the better pilots carried excess airspeed into the merge. This resulted in a larger, slower turn (i.e., decreased turn rate) back into the bogey thus giving the bogey an earlier advantage which was ultimately converted to a win. For V_c (see Figure B-12), significant differences as a function of Outcome occurred only for the Head-on start position. V_c for Draws was less than for either Wins or Losses. Although such differences for the other two start positions were not significant, the trends were the same. These results are in concert with the overall main effect findings that V_c can primarily differentiate Draws.

Two positional advantage measures also produced significant Start Position by Outcome interactions, AAMI Defensive and AAMI Mean > 0. For AAMI Defensive (see Figure B-13), scores were significantly higher for the Head-on start position only for Wins. For AAMI Mean > 0 (see Figure B-14), scores were significantly higher for the Head-on start position only for Losses. As previously discussed, there is an increase in AAMI scores at the beginning of each head-on engagement. It seems likely that such increments due to the initial pass account for these findings. In the case of a Win, the AAMI Defensive score is generally low, so that the initial pass adds significantly to the score. For losses, the AAMI Mean > 0 scores are generally low, so that the initial pass also adds significantly to this score.

Opponent by Outcome. Two aircraft state measures produced a significant Opponent by Outcome interaction, Vertical Velocity and V_c. For Vertical Velocity (see Figure B-15), the results indicated significant differences due to Opponent only for Losses, although the trends for the remaining outcome conditions were the same. The propensity of the AML to fight in the horizontal dimension has already been discussed. These data suggest that for losses against the Human opponent, the subject is forced into a vertical fight and is eventually killed. For losses against the AML, which were quite few, there is even less movement in the vertical dimension and the subjects tend to remain almost completely in the horizontal plane. For V_c (see Figure B-16), the results indicate that the observed main effect of reduced V_c only for Draws is due primarily to those engagements flown against the Human opponent. In fact, there were no differences in V_c as a function of Outcome for engagements flown against the AML. These data also indicate a significantly higher V_c for Wins against the Human opponent when compared against either Draws or Losses. These data further confirm the point made earlier that the Opponent factor not only had a major impact on outcome of the engagement, but also the manner in which the engagement was flown.

Two positional advantage measures also produced a significant Opponent by Result interaction, AAMI Offensive and AAMI Mean > 0. These results are presented in Figures B-17 and B-18. The same pattern of results emerged

from these two measures. Against the AML, a fairly linear relationship with Outcome occurred. Against the Human opponent, however, differences occurred only between Wins and Draws/Losses, but not between Draws and Losses. In fact for the Lose outcome, there were no differences in scores achieved against the AML or the Human opponent. One would intuitively predict that since most engagements were lost against the Human opponent and only a few were lost against the AML, that differences would occur. These results suggest that there may be a ceiling effect operating for the measures obtained under these conditions. Further investigation of these measures is required to clarify these results.

Development of Composite Measures

The final set of analyses attempted to combine the individual candidate measures to produce a composite measure of performance that could be used to reliably predict engagement outcome. An additional question addressed in the analyses was the relative contribution of aircraft state versus positional advantage measures. For all analyses, two models were produced. The first defined Outcome in terms of Win/Draw/Lose while the second defined Outcome in terms of only Win/Lose. Because of the salience of the Opponent factor, separate analyses were performed for each subsample.

The results of these analyses are summarized in Table 10. From these results, it is clear that measures of position advantage are much more predictive of Outcome than the candidate aircraft state measures. Such results are not too surprising, and in fact would be expected. In essence, these results confirm the expectation that being highly offensive is associated with "winning" while conversely, being defensive, is associated with "losing." Such findings, which merely confirm the obvious in some sense, nevertheless do provide evidence for the validity of the AAMI as a measure of positional advantage. Although measures of position advantage are most important, the data also suggest that the addition of state parameters adds to the overall predictiveness of the composite measure, especially for the three-group Outcome model which includes Draws.

It is also clear from the data that the composite measure tends to do a better job of predicting outcome against the AML for the two-group Win/Lose criterion model. The composite measure correctly classified outcome in 97.1% of the cases flown against the AML compared to only 83.5% for engagements flown against the Human opponent. For the three-group Win/Draw/Lose model, approximately 70% of engagements were correctly classified for both opponents.

The importance of Opponent was further confirmed from the results of analyses attempting to predict the results of one opponent from classification functions based upon the other opponent. The results (see Table 12) were quite clear that such functions based upon one opponent do not generalize. Moreover, the results (see Table 13) indicated that classification functions based upon the entire data set, do a poor job of predicting outcome within either of

the two subsamples, despite the fact that the overall classification rates for the entire sample are quite high. The fact that these represent separate distributions is further confirmed by the results (see Tables 13 and 14) of attempts to develop expanded four- and six-group classification functions.

Overall, these results indicate that development of a single composite measure that is predictive of engagement outcome for the entire engagement sample is not possible, at least from the candidate measurement set explored in the present investigation. However, it should be pointed out that, at least for the positional advantage measures, the same measures were selected for each classification function. The measures were the same—only the classification weights differed. In other words, the problem may not be in the candidate measures, per se, but rather on the sample obtained for development of empirical weights necessary to produce a composite measure.

The development of measures for ACM is problematic in that the "goodness" of individual performance is largely based on the skill of the adversary. At the beginning of this investigation, the choice seemed to be either the risk of many adversaries with a wide range of abilities or the risk of limited adversaries that represented the extremes of abilities. The decision to standardize and control the investigation to as large a degree as possible, led to the selection of the latter alternative. In this case, a computer-driven adversary and a single human opponent were chosen as the best way to ensure standardization. Unfortunately, these were found to represent opposite ends of the skill continuum with the result that models developed to predict behavior against one of these opponents do not generalize to the other.

Another problem concerns the criterion, engagement outcome. Wins in the present investigation included not only ordnance kills, but also outcomes such as terrain crashes and over-Gs by the adversary. Moreover, the decision to either "re-fly" the engagement or code the outcome as a "kill" depended upon the time elapsed during the engagement. As previously stated, this represented one of the practical constraints of data collection. The net effect is the introduction of some degree of confounding into the criterion. Certainly, the argument could be made to include only those engagements that ended in an ordnance kill as the subsample on which to attempt the development of a composite measure.

Despite these problems and the inability to produce a single composite measure that predicts well for both samples, these results nonetheless are encouraging in that they do suggest that the development of a composite measure useful for training effectiveness evaluations is indeed feasible. An issue for future investigations is the extent to which such classification functions must predict actual outcomes in order to be considered useful. For example, the current functions correctly classified about 70% of outcomes for the Win/Draw/Lose criterion. Is 70% sufficient? Must we achieve 90%? Moreover, the issue of measurement reliability should be addressed directly. Recall that the rationale for the present effort was to develop a composite measure that is continuously computed throughout the engagement since weapons events

are quite unreliable. Future investigations should assess the improvements resulting from the use of such continuous measures directly.

Recommendations for Future Research

Based on this preliminary analysis of the engagement database, the following recommendations are offered to guide future efforts.

1. Order effects can be ignored in future analyses. The lack of any large, consistent, or interpretable effects due to order of presentation suggests that this factor may be ignored in future analyses of this engagement database.

2. The two line-abreast conditions, i.e., Bogey-Left and Bogey-Right, may be combined to form a single category for future analyses. It is further recommended that an attempt be made to determine whether the effects due to Start Position can be eliminated altogether.

3. Engagements flown against the two adversaries must be analyzed separately. The overwhelming finding from the present study was that Opponent impacted not only Outcome, but also maneuvering as reflected by the aircraft state measures.

4. Of the two sets, the positional advantage measures were most predictive of engagement outcome. This suggests that future efforts should further explore these types of measures. At a minimum, future efforts should also evaluate the usefulness of the Relative Offensive Maneuvering Potential (ROMP) scores which is simply the difference between the AAMI Offensive scores for the two opponents. Furthermore, the weapons range models that are used in calculation of the AAMI should be improved.

5. The aircraft state measures should be further explored since they provide additional information not available from simply the positional advantage measures. Different summary measures, such as the variability of the measure across the engagement, should be evaluated. Different state measures should also be evaluated, with emphasis on those that reflect relative differences between adversaries. Research is also needed that explores whether measurement start-stop logics for specific measures can be developed for ACM. The argument can be made that each of the measures are important only during certain segments of an engagement. Unfortunately, there are no rules whereby such segments are readily defined. Research is required to address this limitation.

6. The potential value of other measures should also be explored. These include such things as measures of energy control, measures of missed opportunities, and measures of maneuvering performance such as lift-vector control.

7. The criterion for engagement outcome should be restricted in future analyses to include only wins and losses due to an ordnance kill.

8. Research is needed to develop a better adversary simulation model. A problem with the current database is that the computer-driven AML was "too easy" while the human opponent was "too difficult." The application of expert systems, including both production systems and artificial neural networks, should be explored.

9. Reliability analyses should be conducted which determine the improvements resulting from use of continuous positional advantage measures when compared against simple outcome event measures.

10. Additional engagement data should be gathered for cross-validation of the measures that prove successful within the current engagement database. We recommend that a round-robin procedure be used in which the experience level of the participants is controlled. We also recommend that engagement data be gathered from the ACMI.

CONCLUSIONS

In view of the stated research objectives for this investigation (see page 8), the following conclusions are offered:

1. Of the four experimental factors investigated, only Opponent significantly affected Outcome. This effect was quite large and is considered to be the most important factor within the engagement database that was gathered. The other factors, Aircraft, Start Position, and Order produced no significant effects.

2. A large number of significant effects were obtained between the experimental factors and the individual candidate measures of performance. The one exception was Order, which produced no significant effects. Aircraft type produced a number of effects upon state measures which were due to differences in the flight control systems of the F-15 and F-16. Start Position also produced effects upon both state and positional advantage measures which were easily explained as a function of the initial set-up conditions. Opponent also produced effects on both types of measures. Differences in the state measures were primarily attributed to "style" differences between the AML and Human opponents, while differences in the position advantage measures were the result of the Human being a much more difficult opponent. And finally, differences occurred in both state and positional advantage measures as a function of engagement Outcome. Strong linear relations were obtained between the "offensive" AAMI measures and Outcome, while the "defensive" AAMI was most discriminative of Losses. These findings provide support for the validity of these measures as predictors of engagement outcome.

3. Of the two categories of measures, aircraft state and positional advantage, the results clearly indicated that the positional advantage type were most highly related to engagement outcome.

4. Composite measures were successfully developed for each Opponent for both a Win/Lose and Win/Draw/Lose criterion. In general, we found that classification accuracy was highest for the Win/Lose criterion. For the Win/Lose/Draw criterion, classification accuracy for both the AML and Human Opponents was roughly 70% while accuracy for the Win/Lose criterion was 84% for the Human and 97% for the AML. Classification functions developed for one opponent (either the AML or Human) did not generalize to the other opponent, despite the fact that the same measures were selected for inclusions into the prediction equation. These findings suggest that the measures, themselves, are valid across opponents, but that development of an optimal set of weights is highly dependent upon the characteristics of the opponent. Based on these results, we concluded that development of a composite measure of ACM performance for use in transfer of training evaluations is certainly feasible.

REFERENCES

Brecke, F.H., & Miller, D.C. (1991). *Aircrew performance measurement in the air combat maneuvering domain: A critical review of the literature.* (AL-TR-1991-0042, AD-B158 404L). Williams AFB, AZ: Armstrong Laboratory, Human Resources Directorate, Aircrew Training Research Division.

Briedenbach, S.T., Ciaviarelli, A.P., Sievers, R., & Lillienthal, M.G. (1985). *Measurement methods and metrics for aircrew assessment during close-in air-to-air combat* (R-85006). San Diego, CA: Cubic Corporation.

Bryce, G.R. (1980). *Data Analysis in RUMMAGE--A User's Guide*. Provo, UT: Brigham Young University.

Ciavaerelli, A.P., Williams, A.M., & Pettigrew, K.W. (1981). *Performance assessment and appraisal system (PAAS)* (NAVTRAEQUIPCEN 78-C-0136-1). Orlando, FL: Naval Training Equipment Center.

Cohen, J. (1965). *Statistical power analysis for the behavioral sciences*. New York, NY: Academic Press.

deBerg, O.H. (1977). *A scoring algorithm for assessing air-to-air combat performance: A factor analysis* (Report No. 05090-77). Maxwell AFB, AL: Air University.

Guilford, J.P. (1965). *Fundamental statistics in psychology and education*. New York, NY: McGraw-Hill.

Jenkins, D.H. (1982). *Simulator training effectiveness evaluation* (TAC Project 79Y-001F). Nellis AFB, NV: USAF Tactical Fighter Weapons Center.

Jennrich, R., & Sampson, P. (1990). Stepwise Discriminant Analysis. In *BMDP Statistical Software Manual Vol 1*, Berkeley, CA: University of California Press.

Kelly, M.J. (1988). Performance measurement during simulated air-to-air combat. *Human Factors, 30*, 495-506.

Kelly, M.J., Wooldridge, L., Hennessy, R.T., Vreuls, D., Barnebey, S.F., Cotton, J.C., & Reed, J.C. (1979). *Air combat maneuvering performance measurement* (AFHRL-TR-79-3/NAVTRAEQUIPCEN IH-35, AD-A077 429). Williams AFB, AZ: Air Force Human Resources Laboratory, Operations Training Division.

Lane, N.E. (1986). *Issues in performance measurement for military aviation with applications to air combat maneuvering.* Orlando, FL: Naval Training Systems Center.

McGuiness, J., Forbes, J.M., & Rhoads, J.E. (1984). *Air combat maneuvering performance measurement system design* (AFHRL-TP-83-56, AD-B081 060L). Williams AFB, AZ: Air Force Human Resources Laboratory, Operations Training Division.

Moran, M.A., Engelman, L., Fitzgerald, G., & Lynch, B. (1990). Polychotomous stepwise logistic regression. In *BMDP Statistical Software Manual Vol 2* Berkeley, CA: University of California Press.

Nunnaly, J.C. (1967). *Psychometric theory.* New York, NY: McGraw-Hill.

Oberle, R.A. (1974). *An air combat maneuver conversion model* (Report No. CRC 274). Washington, DC: Office of Naval Research.

Payne, T.A., Hirsch, D.L., Semple, C.A., Farmer, J.R., Spring, W.G., Sanders, M.S., Wimer, C.A., Carter, V.E., and Hu, A. (1976). *Experiments to evaluate advanced flight simulation in air combat pilot training. Vol 1. Transfer of learning experiment.* Hawthorne, CA: Northrop Corporation.

Pohlmann, L.D., & Reed, J.C. (1978). *Air-to-air combat skills: Contribution of platform motion to initial training* (AFHRL-TR-78-53, AD-A062 738). Williams AFB, AZ: Air Force Human Resources Laboratory, Flying Training Division.

Simpson, W.R. (1976). *Development of a time-variant figure-of-merit for use in analysis of air combat maneuvering engagements* (TM-76-1 SA). Patuxent River, MD: Naval Air Test Center.

Stacy, W.E. (1980). *Project progress to Fleet Analysis Center for ACMR Readiness Estimation System* (Contract No. N00123-79-0807). Corona, CA: Value Systems Engineering.

Waag, W.L., Pierce, B.J., & Fessler, S. (1987). *Performance measurement requirements for tactical aircrew training* (AFHRL-TR-86-62, AD-A183 814). Williams AFB, AZ: Air Force Human Resources Laboratory, Operations Training Division.

APPENDIX A
DESCRIPTIVE STATISTICS

DESCRIPTIVE STATISTICS FOR AIRCRAFT STATE MEASURES

VARIABLE NAME	GROUPING VARIABLE	TOTAL FREQ.	MEAN	STANDARD DEVIATION	SMALLEST VALUE	SMALLEST Z-SC	LARGEST VALUE	LARGEST Z-SC
VERTICAL		1000	66.545	78.946	-310.76	-4.78	306.51	3.04
VELOCITY	OUTCOME							
	WIN	456	49.090	88.921	-250.15	-3.37	306.51	2.89
	DRAW	150	46.597	32.923	-51.601	-2.98	116.17	2.11
	LOSE	394	94.340	70.692	-310.76	-5.73	265.26	2.42
	OPPON							
	AML	500	37.626	82.799	-250.15	-3.48	306.51	3.25
	HUMAN	500	95.463	62.815	-310.76	-6.47	265.26	2.70
	VERSION							
	ORDER_1	248	54.771	77.091	-206.92	-3.39	284.72	2.98
	ORDER_2	256	71.166	75.938	-216.76	-3.79	306.51	3.10
	ORDER_3	256	62.355	83.664	-310.76	-4.46	270.33	2.49
	ORDER_4	240	78.250	77.234	-192.86	-3.51	283.13	2.65
	AIRCRAFT							
	F-15	480	64.053	85.884	-310.76	-4.36	270.33	2.40
	F-16	520	68.844	71.957	-216.76	-3.97	306.51	3.30
	STARTP							
	HEAD_ON	501	59.714	84.495	-310.76	-4.38	306.51	2.92
	BOGEY_R	250	72.094	73.221	-190.71	-3.59	290.31	2.98
	BOGEY_L	249	74.717	71.668	-139.09	-2.98	284.72	2.93
ROLL RATE		1000	-.00815	.07468	-.32300	-4.22	.31500	4.33
	OUTCOME							
	WIN	456	-.00294	.07906	-.32300	-4.05	.31500	4.02
	DRAW	150	-.00494	.04563	-.15800	-3.35	.15000	3.40
	LOSE	394	-.01540	.07787	-.30800	-3.76	.25100	3.42
	OPPON							
	AML	500	-.00521	.07214	-.32300	-4.41	.31500	4.44
	HUMAN	500	-.01109	.07710	-.30800	-3.85	.28000	3.78

VARIABLE NAME	GROUPING VARIABLE	TOTAL FREQ.	MEAN	STANDARD DEVIATION	SMALLEST VALUE	SMALLEST Z-SC	LARGEST VALUE	LARGEST Z-SC
	VERSION							
	ORDER_1	248	-.00454	.06896	-.25692	-3.66	.25100	3.71
	ORDER_2	256	-.01061	.07704	-.30800	-3.86	.24500	3.32
	ORDER_3	256	-.00611	.07204	-.32300	-4.40	.28000	3.97
	ORDER_4	240	-.01143	.08057	-.20600	-2.41	.31500	4.05
	AIRCRAFT							
	F-15	480	-.00804	.07821	-.32300	-4.03	.31500	4.13
	F-16	520	-.00825	.07134	-.30800	-4.20	.25100	3.63
	STARTP							
	HEAD_ON	501	-.00798	.06907	-.30800	-4.34	.24500	3.66
	BOGEY_R	250	.01445	.07780	-.19900	-2.74	.31500	3.86
	BOGEY_L	249	-.03119	.07564	-.32300	-3.86	.18700	2.88
TURN RATE		1000	.22821	.03978	.11800	-2.77	.35860	3.28
	OUTCOME							
	WIN	456	.20775	.03203	.12200	-2.68	.32300	3.60
	DRAW	150	.23351	.03304	.16400	-2.10	.33400	3.04
	LOSE	394	.24987	.03812	.11800	-3.46	.35860	2.85
	OPPON							
	AML	500	.20364	.02754	.12200	-2.96	.28200	2.84
	HUMAN	500	.25278	.03463	.11800	-3.89	.35860	3.06
	VERSION							
	ORDER_1	248	.22498	.04038	.12200	-2.55	.33400	2.70
	ORDER_2	256	.23170	.03994	.13600	-2.40	.35860	3.18
	ORDER_3	256	.22548	.03840	.11800	-2.80	.33800	2.93
	ORDER_4	240	.23075	.04018	.12400	-2.66	.35500	3.09
	AIRCRAFT							
	F-15	480	.22245	.03802	.11800	-2.75	.33800	3.04
	F-16	520	.23352	.04065	.12200	-2.74	.35860	3.08

VARIABLE NAME	GROUPING VARIABLE	TOTAL FREQ.	MEAN	STANDARD DEVIATION	SMALLEST VALUE	SMALLEST Z-SC	LARGEST VALUE	LARGEST Z-SC
STARTP								
	HEAD_ON	501	.22015	.03876	.11800	-2.64	.33900	3.07
	BOGEY_R	250	.23961	.04023	.12400	-2.87	.35500	2.87
	BOGEY_L	249	.23300	.03787	.12200	-2.93	.35860	3.32
G-LOADING		1000	3.5324	.86584	1.6770	-2.14	7.4120	4.48
	OUTCOME							
	WIN	456	3.4369	.86470	1.6770	-2.04	6.2980	3.31
	DRAW	150	3.4205	.87060	1.7850	-1.88	7.4120	4.58
	LOSE	394	3.6854	.84488	2.0660	-1.92	6.6660	3.53
	OPPON							
	AML	500	3.3695	.84699	1.6770	-2.00	7.4120	4.77
	HUMAN	500	3.6952	.85459	1.7850	-2.24	6.7520	3.58
	VERSION							
	ORDER_1	248	3.4856	.89075	1.6770	-2.03	6.6660	3.57
	ORDER_2	256	3.6446	.85768	2.1050	-1.80	6.7520	3.62
	ORDER_3	256	3.5027	.89936	1.8640	-1.82	7.4120	4.35
	ORDER_4	240	3.4925	.80463	1.7850	-2.12	6.1170	3.26
	AIRCRAFT							
	F-15	480	3.7131	.85951	1.7850	-2.24	6.2980	3.01
	F-16	520	3.3655	.83849	1.6770	-2.01	7.4120	4.83
	STARTP							
	HEAD_ON	501	3.5318	.85375	1.8090	-2.02	7.4120	4.54
	BOGEY_R	250	3.5587	.87230	1.8370	-1.97	6.1170	2.93
	BOGEY_L	249	3.50[illegible]	.88599	1.6770	-2.07	6.2980	3.15
AIRSPEED		1000	291.90	68.155	147.88	-2.11	575.09	4.16
	OUTCOME							
	WIN	456	303.80	69.043	159.69	-2.09	575.09	3.93
	DRAW	150	273.31	61.320	147.88	-2.05	548.13	4.48
	LOSE	394	285.19	67.243	173.74	-1.66	535.56	3.72

VARIABLE NAME	GROUPING VARIABLE	TOTAL FREQ.	MEAN	STANDARD DEVIATION	SMALLEST VALUE	SMALLEST Z-SC	LARGEST VALUE	LARGEST Z-SC
	OPPON							
	AML	500	302.98	69.262	159.69	-2.07	575.09	3.93
	HUMAN	500	280.81	65.238	147.88	-2.04	535.56	3.90
	VERSION							
	ORDER_1	248	288.33	67.777	159.69	-1.90	535.56	3.65
	ORDER_2	256	297.70	68.336	185.91	-1.64	575.09	4.06
	ORDER_3	256	290.88	67.144	173.74	-1.74	548.13	3.83
	ORDER_4	240	290.47	69.452	147.88	-2.05	508.40	3.14
	AIRCRAFT							
	F-15	480	306.70	65.754	147.88	-2.42	516.63	3.19
	F-16	520	278.23	67.535	159.69	-1.76	575.09	4.40
	STARTP							
	HEAD_ON	501	305.14	71.268	186.40	-1.67	575.09	3.79
	BOGEY_R	250	279.87	61.274	159.69	-1.96	458.37	2.91
	BOGEY_L	249	277.32	63.150	147.88	-2.05	498.80	3.51
CLOSING		1000	69.885	95.105	-287.68	-3.76	524.55	4.78
VELOCITY	OUTCOME							
	WIN	456	47.318	98.503	-287.68	-3.40	461.51	4.20
	DRAW	150	28.992	28.529	-74.464	-3.61	131.36	3.59
	LOSE	394	111.57	92.081	-35.056	-1.59	524.55	4.49
	OPPON							
	AML	500	31.590	76.966	-287.68	-4.15	353.96	4.19
	HUMAN	500	108.18	96.155	-35.056	-1.49	524.55	4.33
	VERSION							
	ORDER_1	248	66.948	82.396	-155.47	-2.70	402.72	4.08
	ORDER_2	256	68.043	97.526	-152.20	-2.26	524.55	4.68
	ORDER_3	256	69.935	98.728	-157.73	-2.31	461.51	3.97
	ORDER_4	240	74.829	101.03	-287.68	-3.59	427.99	3.50

VARIABLE NAME	GROUPING VARIABLE	TOTAL FREQ.	MEAN	STANDARD DEVIATION	SMALLEST VALUE	SMALLEST Z-SC	LARGEST VALUE	LARGEST Z-SC
	AIRCRAFT							
	F-15	480	74.940	92.025	-157.73	-2.53	477.49	4.37
	F-16	520	65.218	97.719	-287.68	-3.61	524.55	4.70
	STARTP							
	HEAD_ON	501	113.61	105.60	-287.68	-3.80	524.55	3.89
	BOGEY_R	250	27.040	54.722	-153.90	-3.31	210.89	3.36
	BOGEY_L	249	24.921	56.611	-145.06	-3.00	266.73	4.27

DESCRIPTIVE STATISTICS FOR POSITIONAL ADVANTAGE MEASURES

VARIABLE NAME	GROUPING VARIABLE	TOTAL FREQ.	MEAN	STANDARD DEVIATION	SMALLEST VALUE	SMALLEST Z-SC	LARGEST VALUE	LARGEST Z-SC
AAMI		1000	33.973	18.816	0.0000	-1.81	80.543	2.48
OFFENSIVE	OUTCOME							
	WIN	456	47.602	15.279	5.2410	-2.77	80.543	2.16
	DRAW	150	28.160	16.119	2.1379	-1.61	72.148	2.73
	LOSE	394	20.411	10.880	0.0000	-1.88	57.000	3.36
	OPPON							
	AML	500	45.572	16.924	2.4935	-2.55	80.543	2.07
	HUMAN	500	22.374	22.360	0.0000	-1.81	67.625	3.66
	VERSION							
	ORDER_1	248	33.555	18.656	0.0000	-1.80	75.403	2.24
	ORDER_2	256	31.725	18.478	.09200	-1.71	77.625	2.48
	ORDER_3	256	34.675	18.859	0.0000	-1.84	74.344	2.10
	ORDER_4	240	36.052	19.130	0.0000	-1.88	80.543	2.33
	AIRCRAFT							
	F15	480	34.503	19.285	.34200	-1.77	78.189	2.27
	F-16	520	33.483	18.378	0.0000	-1.82	80.543	2.56
	STARTP							
	HEAD_ON	501	36.644	17.945	2.7830	-1.89	77.625	2.28
	BOGEY_R	250	31.504	19.650	0.0000	-1.60	80.543	2.50
	BOGEY_L	249	31.076	18.981	0.0000	-1.64	75.403	2.34
AAMI		1000	28.707	19.417	0.0000	-1.48	85.312	2.92
DEFENSIVE	OUTCOME							
	WIN	456	14.785	11.217	0.0000	-1.32	60.263	4.05
	DRAW	150	23.963	16.045	3.2780	-1.29	80.201	3.51
	LOSE	394	46.627	12.638	5.2420	-3.27	85.312	3.06
	OPPON							
	AML	500	12.280	7.4791	0.0000	-1.64	51.677	5.27
	HUMAN	500	45.135	12.573	2.6920	-3.38	85.312	3.20

VARIABLE NAME	GROUPING VARIABLE	TOTAL FREQ.	MEAN	STANDARD DEVIATION	SMALLEST		LARGEST	
					VALUE	Z-SC	VALUE	Z-SC
	VERSION							
	ORDER_1	248	27.699	18.991	0.0000	-1.46	78.000	2.65
	ORDER_2	256	30.164	19.319	2.7340	-1.42	80.201	2.59
	ORDER_3	256	29.155	19.896	2.6920	-1.33	85.312	2.82
	ORDER_4	240	27.718	19.445	2.3917	-1.30	83.253	2.83
	AIRCRAFT							
	F-15	480	28.424	18.697	0.0000	-1.52	83.253	2.93
	F-16	520	28.969	20.074	0.0000	-1.44	85.312	2.81
	STARTP							
	HEAD_ON	501	30.929	18.984	3.5500	-1.44	80.201	2.60
	BOGEY_R	250	26.492	19.701	0.0000	-1.34	83.253	2.88
	BOGEY_L	249	26.461	19.557	0.0000	-1.35	85.312	3.01
AAMI % TIME		1000	55.187	18.588	0.0000	-2.97	97.235	2.26
OVER ZERO	OUTCOME							
	WIN	456	67.195	14.555	21.159	-3.16	97.235	2.06
	DRAW	150	51.576	17.859	6.3190	-2.53	93.681	2.36
	LOSE	394	42.665	13.514	0.0000	-3.16	76.220	2.48
	OPPON							
	AML	500	64.820	16.934	9.3150	-3.28	97.235	1.91
	HUMAN	500	45.555	14.801	0.0000	-3.08	89.000	2.94
	VERSION							
	ORDER_1	248	55.866	18.446	0.0000	-3.03	95.322	2.14
	ORDER_2	256	52.220	18.074	2.6140	-2.74	94.805	2.36
	ORDER_3	256	55.241	18.336	0.0000	-3.01	93.681	2.10
	ORDER_4	240	57.594	19.227	0.0000	-3.00	97.235	2.06
	AIRCRAFT							
	F-15	480	55.327	18.741	4.7950	-2.70	94.805	2.11
	F-16	520	55.059	18.463	0.0000	-2.98	97.235	2.28

VARIABLE NAME	GROUPING VARIABLE	TOTAL FREQ.	MEAN	STANDARD DEVIATION	SMALLEST VALUE	SMALLEST Z-SC	LARGEST VALUE	LARGEST Z-SC
	STARTP							
	HEAD_ON	501	56.425	17.595	6.3190	-2.85	94.805	2.18
	BOGEY_R	250	54.333	19.733	0.0000	-2.75	93.909	2.01
	BOGEY_L	249	53.556	19.242	0.0000	-2.78	97.235	2.27
AAMI MEAN		1000	57.153	18.078	0.0000	-3.16	93.845	2.03
OVER ZERO	OUTCOME							
	WIN	456	69.434	12.231	24.767	-3.65	93.845	2.00
	DRAW	150	51.104	15.590	14.585	-2.34	84.000	2.11
	LOSE	394	45.242	15.305	0.0000	-2.96	91.200	3.00
	OPPON							
	AML	500	68.070	13.210	14.585	-4.05	93.845	1.95
	HUMAN	500	46.236	15.524	0.0000	-2.98	91.200	2.90
	VERSION							
	ORDER_1	248	55.831	18.272	0.0000	-3.06	93.845	2.08
	ORDER_2	256	55.977	17.873	3.5000	-2.94	91.200	1.97
	ORDER_3	256	58.348	18.315	0.0000	-3.19	90.000	1.73
	ORDER_4	240	58.499	17.768	0.0000	-3.29	89.766	1.76
	AIRCRAFT							
	F-15	480	57.960	18.137	7.[illegible]430	-2.80	92.333	1.90
	F-16	520	56.408	18.009	0.0000	-3.13	93.845	2.08
	STARTP							
	HEAD_ON	501	61.741	15.372	26.861	-2.27	93.845	2.09
	BOGEY_R	250	52.510	19.735	0.0000	-2.66	90.265	1.91
	BOGEY_L	249	52.583	19.075	0.0000	-2.76	89.904	1.96

APPENDIX B
SIGNIFICANT INTERACTION EFFECTS

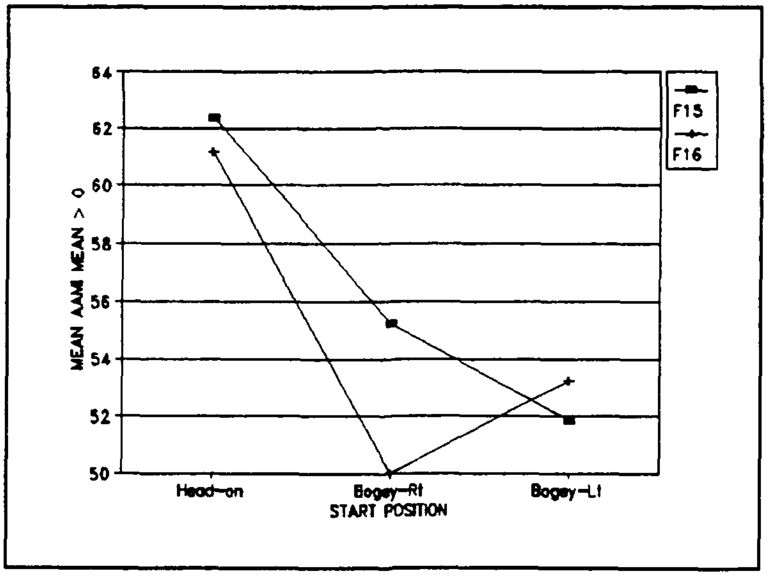

Figure B-3. AAMI Mean > 0 for Aircraft by Start Position.

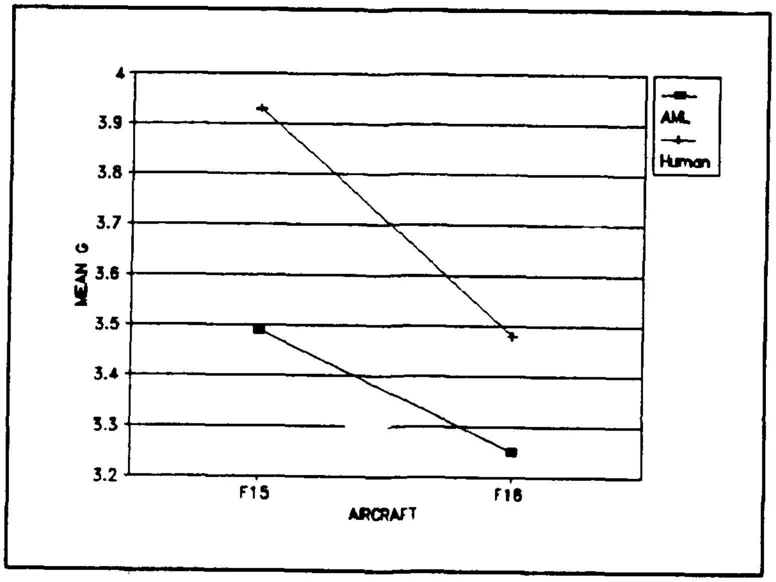

Figure B-4. G for Aircraft by Opponent.

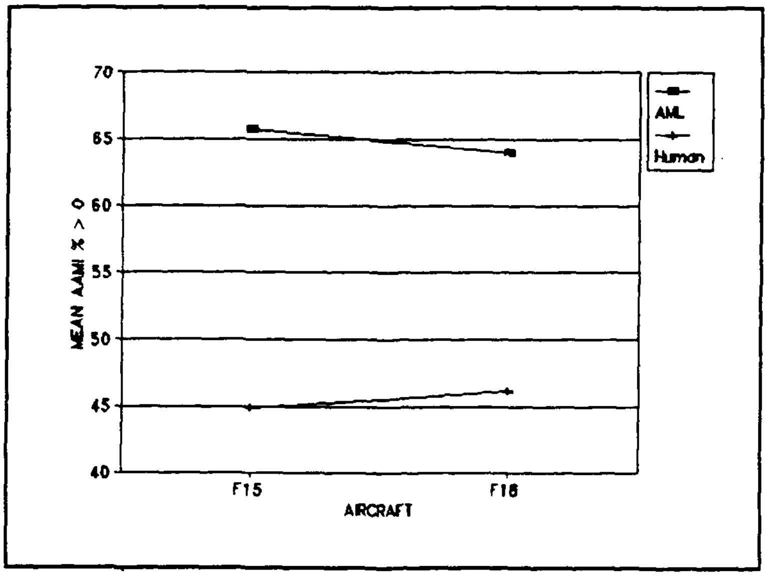

Figure B-5. AAMI % > 0 for Aircraft by Opponent.

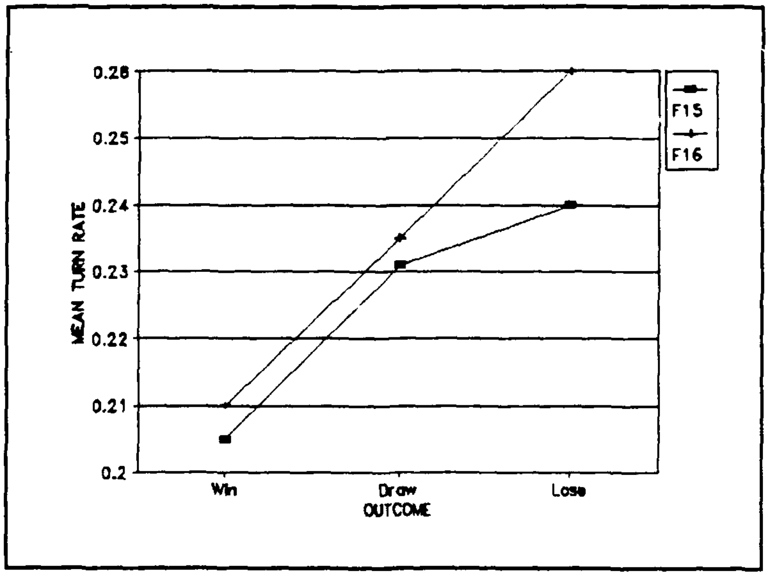

Figure B-6. Turn Rate for Aircraft by Outcome.

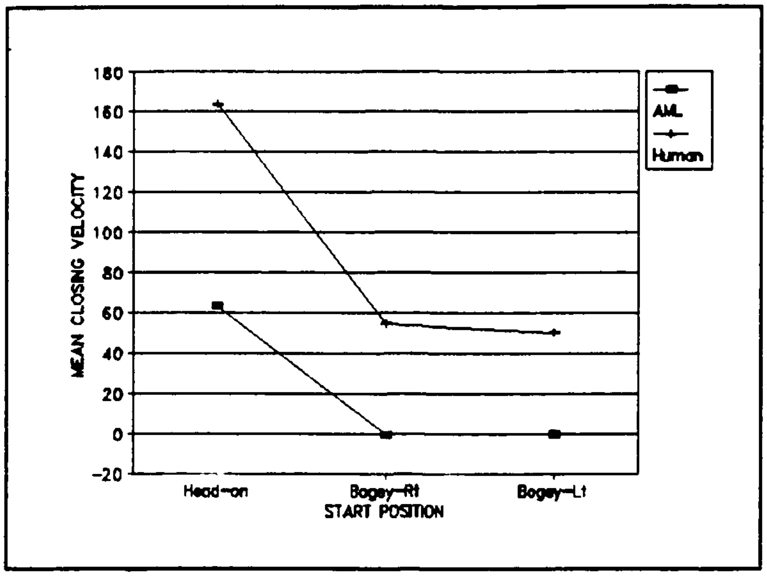

Figure B-7. Closing Velocity for Start Position by Opponent.

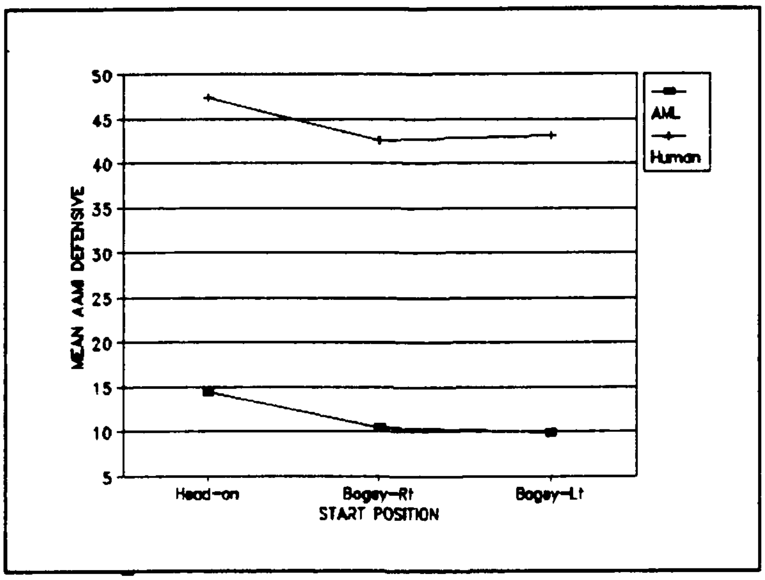

Figure B-8. AAMI Defensive for Start Position by Opponent.

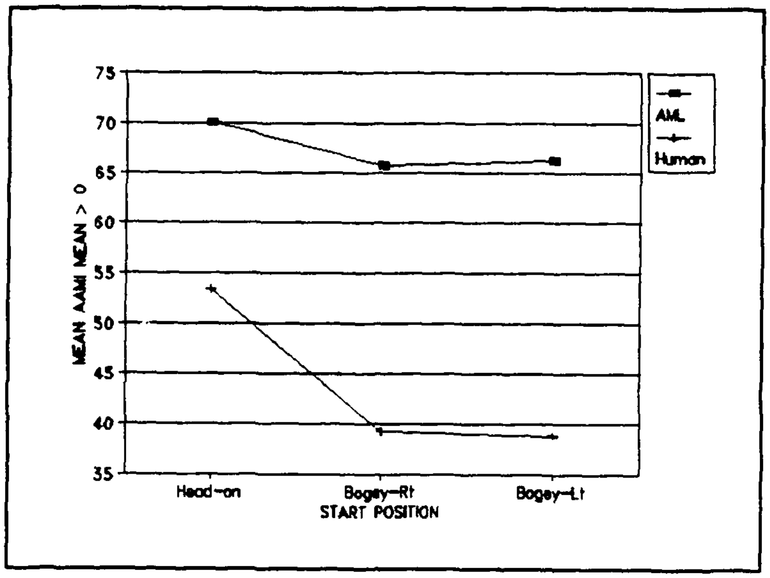

Figure B-9. AAMI Mean > 0 for Start Position by Opponent.

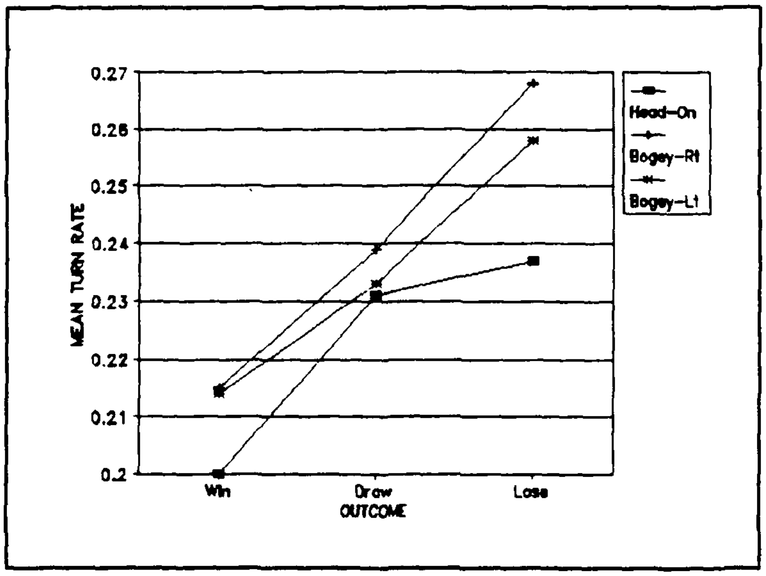

Figure B-10. Turn Rate for Start Position by Outcome.

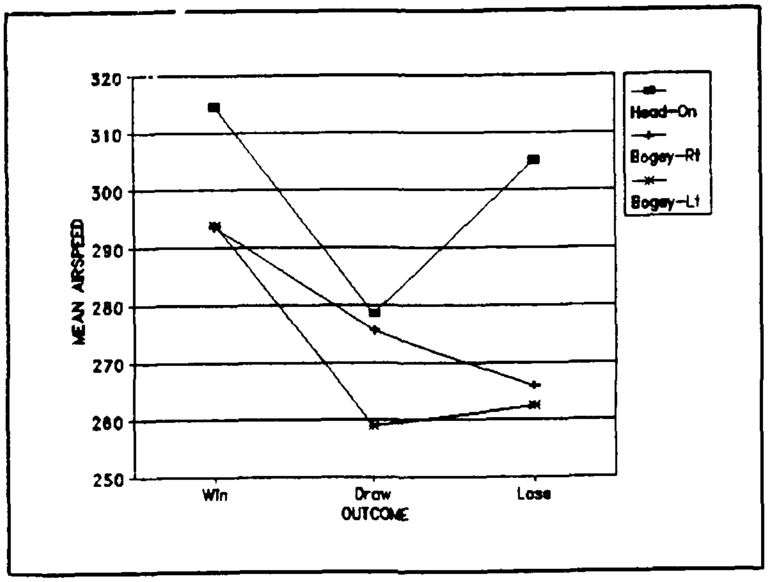

Figure B-11. Airspeed for Start Position by Outcome.

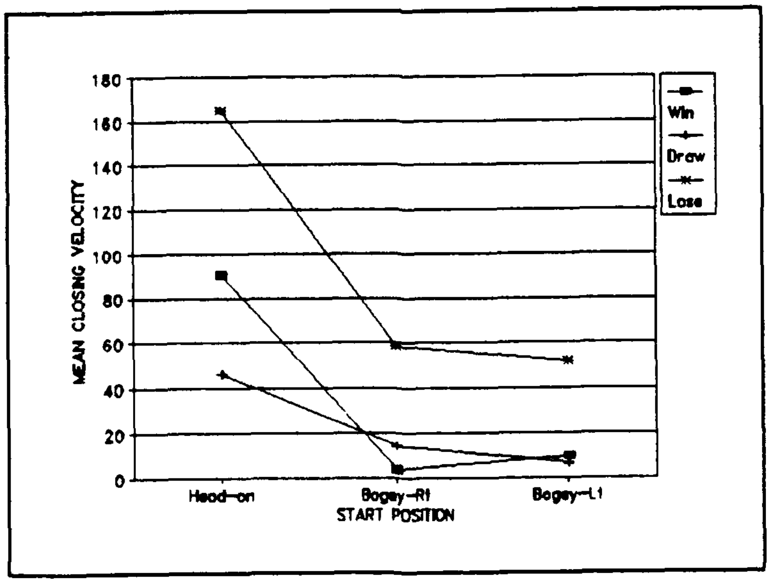

Figure B-12. Closing Velocity for Start Position by Outcome.

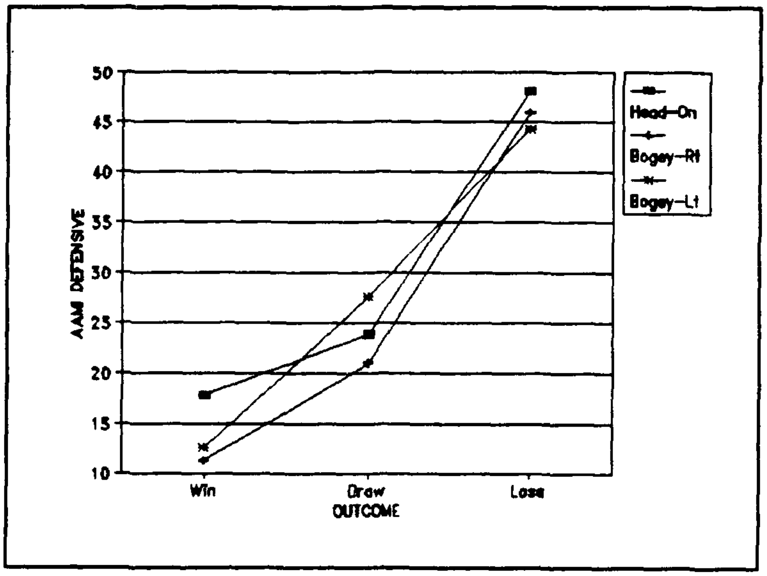

Figure B-13. AAMI Defensive for Start Position by Outcome.

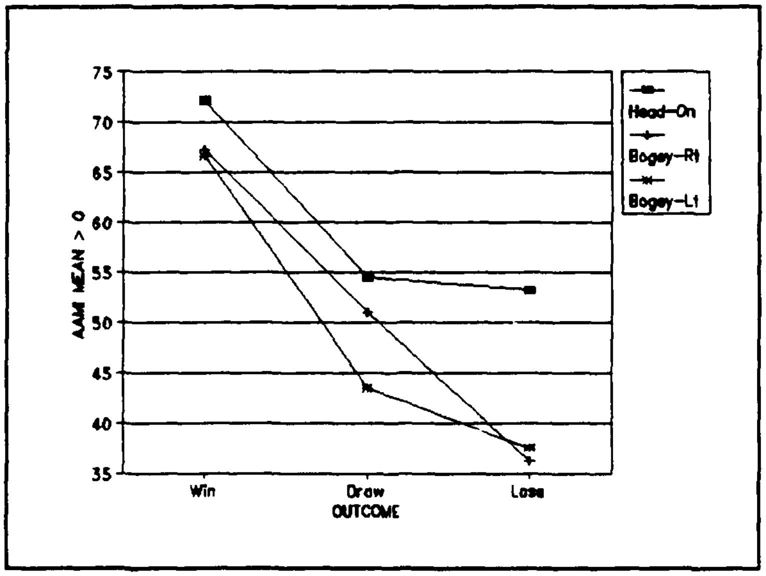

Figure B-14. AAMI Mean > 0 for Start Position by Outcome.

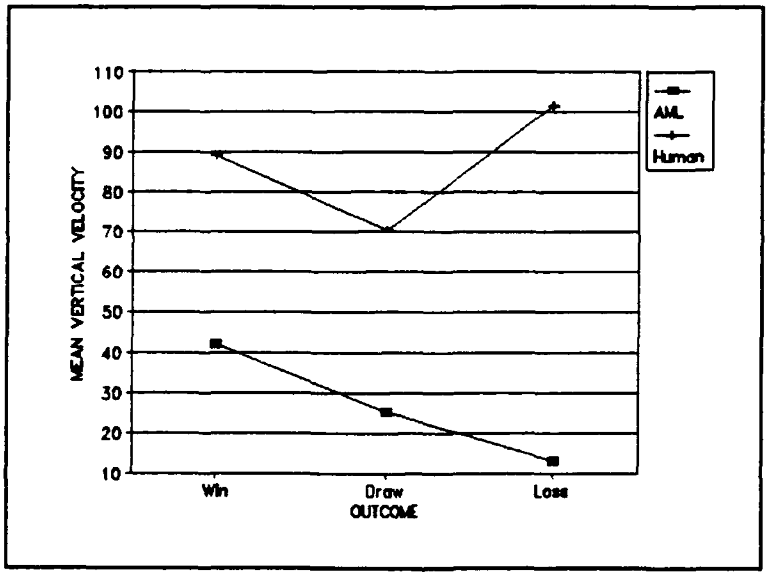

Figure B-15. Vertical Velocity for Outcome by Opponent.

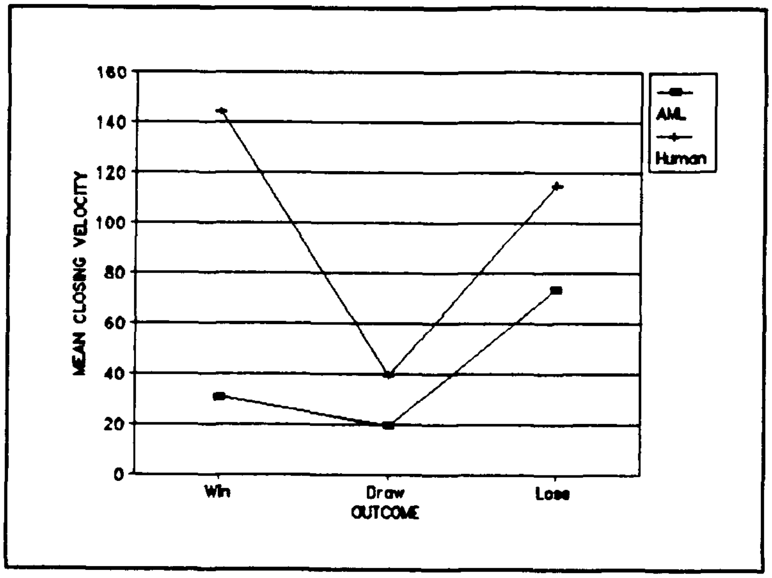

Figure B-16. Closing Velocity for Outcome by Opponent.

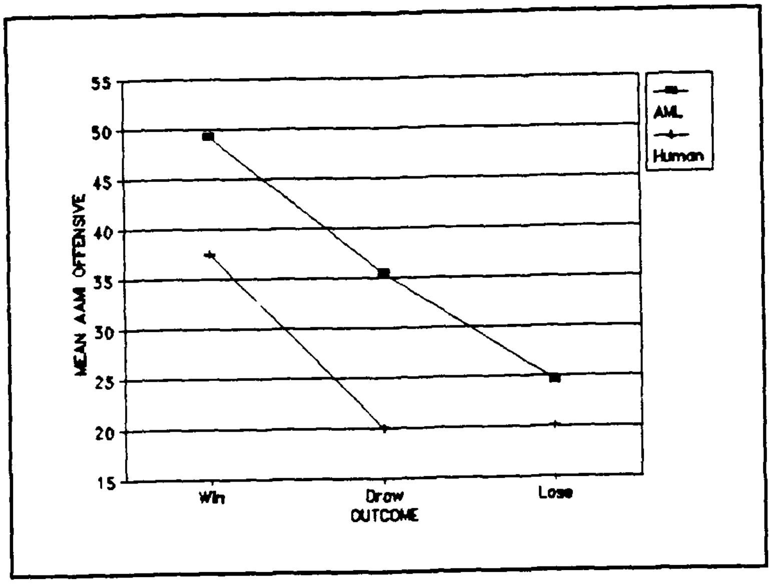

Figure B-17. AAMI Offensive for Outcome by Opponent.

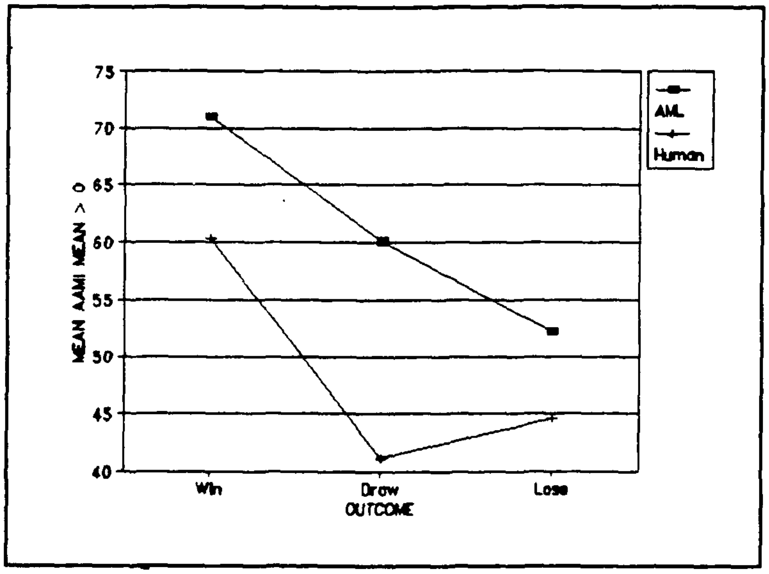

Figure B-18. AAMI Mean > 0 for Outcome by Opponent.

*U S GOVERNMENT PRINTING OFFICE 1992-0-661-049 60021